

The British Journal for the Philosophy of Science

VOLUME I

FEBRUARY, 1951

No. 4

THE NATURE OF SOME OF OUR PHYSICAL CONCEPTS

I ★

IN these lectures we shall try to develop an awareness of the operational content of several of the important concepts of physics. The profitableness of such an operational awareness was perhaps first forced on the attention of physicists by the special theory of relativity of Einstein and later by quantum mechanics. I believe that there are still unexplored implications in the macroscopic concepts of classical physics, and that our understanding of some of these concepts and our method of handling them is not even yet satisfactory in all respects. We shall attempt here to study some of the implications of pushing our operational analysis further than is customary.

The fundamental idea back of an operational analysis is simple enough; namely that we do not know the meaning of a concept unless we can specify the operations which were used by us or our neighbour in applying the concept in any concrete situation. If we can thus specify the operations we can to a large extent reproduce the situation which was encountered by us or our neighbour and which we are trying to communicate. The operational aspect is not by any means the only aspect of meaning, but it is often the most important single aspect, particularly in scientific situations and in a society with as homogeneous a background as ours. In making an operational analysis we are dealing with necessary, as distinguished from sufficient,

* The first of three lectures delivered in the University of London (University College) on 24th, 26th, 28th April 1950.

conditions. I have already expounded in several writings¹ various consequences of the operational point of view, so that only a comparatively brief indication of what is involved is necessary here. In this analysis the concept of operation is itself accepted as unanalysed. It is a matter of experience that we can perform certain operations at will, and that our neighbour can perform the 'same' operations. Operations have a certain repeatability and identifiability and therefore 'objectivity'; this we take as a matter of observation and do not inquire how it comes about that this is true.

The operations which are important in the formation of the concepts of physics, or of science in general, may be of different kinds. In the first place there are the operations of the laboratory, or instrumental operations, in many cases operations of measurement. The sense organs are here to be considered as instruments. It was in an analysis of the instrumental operations of measuring lengths and times that Einstein discovered those overlooked features that are the basis of special relativity, and the present operational attitude of many physicists toward the concepts of physics largely stems from this analysis of Einstein. It is often supposed that the operational criterion of meaning demands that the operations which give meaning to a physical concept *must* be instrumental operations. This is, I believe, palpably a mistaken point of view, for simple observation shows that physicists do profitably employ concepts the meaning of which is not to be found in the instrumental operations of the laboratory, and which cannot be reduced to such operations without residue. Nearly all the concepts of theoretical or mathematical physics are of this character, such for example as the stress inside an elastic body subject to surface forces, or the ψ function of wave mechanics. In fact, there is hardly any physical concept which does not enter to a certain extent into some theoretical edifice and which does not therefore possess to a certain degree a non-instrumental component. All these non-instrumental operations we may loosely lump together as 'mental' operations. Among the many mental operations we may single out for special attention the sort of operations performed by the theoretical physicist in his mathematical manipulations and characterise these as 'paper-and-pencil' operations. Among the paper-and-pencil operations are to be included all manipulations with symbols, whether or not the symbols are the conventional symbols of mathematics.

¹ *The Logic of Modern Physics*, New York, 1927; *The Nature of Physical Theory*, Princeton, 1936; several chapters in *Reflections of a Physicist*, New York, 1950.

SOME OF OUR PHYSICAL CONCEPTS—I

It will usually be sufficient to recognise only these two kinds of operations, namely instrumental and paper-and-pencil operations. Sometimes, however, it will be profitable to recognise other sorts of mental operation than the pencil-and-paper operation, although the line of separation is by no means sharp. Probably the most important of these are the verbal operations. Civilised man lives to a large extent in a verbal world of his own making: in this verbal world he exhibits patterns of behaviour which he finds no less compelling than the patterns forced on him by the 'external' objects of the physical world. He can make verbal experiments, as by asking himself 'would I say thus and thus in such and such a situation?' The sort of physical concept which he finds it profitable or at least congenial to use is usually determined to a certain extent by his verbal demands as disclosed by such verbal experiments. The formal element or the element of pure construction, in our physical concepts is often determined by such verbal demands. In our analysis we shall recognise several situations in which the demands of our verbalisation have been effectively directive.

Not only do we make verbal experiments, but we also make experiments in connection with the paper-and-pencil component of our operations. These are often referred to as 'mental' experiments, without other qualification, although strictly a verbal experiment must also be recognised as 'mental.' The conventional mental experiment of the physicist is a highly idealised experiment with conceptualised physical instruments, ignoring in many cases physical limitations which in principle would make the experiment impossible. For example, in the classical electron theory of Lorentz, a meaning was given to the electrostatic force at points inside the electron in terms of the same conceptual operations which give meaning to the ordinary macroscopic electrostatic field, ignoring the essential physical fact that no charge smaller than an electron exists to serve as the test body in determining and giving meaning to the field. It will be found, I think, that in many of the situations of theoretical physics our meanings are to be sought in terms of mental experiments of this sort. This is, however, by no means the case in all the situations of theoretical physics, but particularly in the abstract situations of wave mechanics we are content to do without the visualisation afforded by the mental experiment and utilise only the paper-and-pencil operations of mathematical symbolisation.

The paper-and-pencil world is a world in which free invention is

possible, divorced from any immediate contact with the instrumental world of the laboratory. In this world of free invention discoveries are possible just as in the laboratory. The matter of nomenclature is not important here—we discover that we can invent. In this world we demand of ourselves that we be able to answer any question in which we can discover a meaning and that we be able to foretell the outcome of any describable mental experiment. Discovering a meaning in questions asked by analogy is one of the methods of discovery, for meanings so discovered may then suggest new experiments in the instrumental world.

It will be seen that a very great latitude is allowed to the verbal and the paper-and-pencil operation. I think, however, that physicists are agreed in imposing one restriction on the freedom of such operations, namely that such operations must be capable of eventually, although perhaps indirectly, making connection with instrumental operations. Only in this way can the physicist keep his feet on the ground or achieve a satisfactory degree of precision; instrumental contact affords the only 'reality' which he accepts as pertinent for him. Doubtless concepts which are not capable of eventual instrumental emergence but which are permanently confined to the verbal domain, like many of those used in daily life, are of the greatest importance in affecting human behaviour. Politics, philosophy and religion are full of such purely verbal concepts; it is merely that such concepts are outside the field of the physicist.

The most important of our paper-and-pencil operations, as already suggested, are doubtless those which we perform when we mathematicise. There are two points to be considered here: the nature of the mathematical operations themselves, and the nature of the step by which we transport ourselves from the universe of instrumental operations to the universe of the operations of mathematics. It is in the first place obvious that mathematical operations are by nature different, in at least one important respect, from the instrumental operations of the laboratory. The latter are always subject to a certain haziness or margin of error, as when we try to push our readings to the limit by estimating the fraction of the smallest division of our instrument. There is no such haziness in mathematics, but any number may be written out to an unlimited number of decimal places (by repetition of zeroes in any event), far beyond the possible precision of any physical measurement. Furthermore, the numbers thus specified by mathematics yield by mathematical manipulation other numbers also

SOME OF OUR PHYSICAL CONCEPTS—I

perfectly sharp, subject to no instrumental or psychological uncertainty, and corresponding to nothing we encounter in physical experience. For instance, the statement of geometry that the length of a straight line connecting two points is less than that of any other joining line, no matter how little it may deviate from the straight line, corresponds to no statement that we are in a position to verify about our actual experience. Such sharpness does not reproduce the structure of experience, and in this respect mathematics fails, if indeed it is the object of mathematics to reproduce this aspect of experience. But it is often stated that the world around us is essentially mathematical in nature and is controlled by laws of mathematical precision. It seems to me that this view cannot be maintained, and that the correspondence with the world of mathematics is not a complete correspondence. When dealing with mathematics we are in a different world from the world of the laboratory. In mathematics we are dealing with paper-and-pencil operations which are palpably our own invention. Although mathematics is an invention, it is obviously a good invention, for only by means of it have we been able to acquire the degree of understanding and control of nature which we now enjoy.

In applying the invention of mathematics to the world about us we are forced to make a logical jump, for which I think there can be no rigorous justification. When we theorise about the results of a measurement, we replace the results of the measurements, with all their inevitable haziness, with mathematically sharp numbers. Just which one of the infinitely many mathematical numbers which would correspond to the measurement within its margin of error we select is entirely undetermined by any logical criterion. The choice is made by some quite extraneous consideration, such as simplicity or ease of manipulation. We do not ordinarily regard it as necessary to justify our choice, and fortify ourselves in our indifference by our extensive past experience which has shown it to be a matter of little moment.

Our discussion up to this point has been on a level where we can treat mathematics and logic as sharp. But it is possible to push the analysis to another level at which their sharpness disappears, the level at which we see mathematics and logic as human activities subject to the haziness of all such enterprise and containing an experimental component. It is not possible to ply the enterprise of logic without the concepts of identity and recurrence—we must be able to assure ourselves that we are making the *same* proposition *twice*, and in giving ourselves this assurance the haze enters.

If mathematics and logic are not perfectly sharp and contain an experimental component, what validity in the instrumental world does a conclusion have which has been reached by mathematical or logical analysis? Is it superfluous to verify a conclusion reached in this way? To make the question specific, suppose that we have established by experiment that the equation $\frac{d^2s}{dt^2} = g$ describes the motion of any

freely falling body. We then deduce from this equation by the logical steps of mathematics that the distance fallen by any body starting from rest is connected with the time of fall by the equation $s = \frac{1}{2}gt^2$. Is it superfluous to verify this result by measurement? I think a common answer could be that verification is quite superfluous, on the ground that the integrated form of the equation is already implied (whatever 'implied' may mean) in the differential equation from which we started. I think a better answer would be that the answer depends on the physical operations which we have chosen to employ for our verification, first of the differential equation, and then of the integrated form. If the operations for these two verifications are such that both verifications may be made from the same set of readings in a note book, then we may say that verification of the integrated equation is superfluous except in so far as it guards against our own blunders. But if the set of readings made to verify the differential equation is not the same as made to verify the integrated equation, then verification is not superfluous. The experimental situation may be treated from either point of view. We might, for example, verify the differential equation from closely spaced readings of position and time, recorded perhaps with a high speed movie camera, the position being given by the readings on a long continuously graduated scale over which the body falls, and the time by the position of the hands of a clock carried with the falling body. In this case the identical set of readings would be used to verify the two equations, and verification may be said to be superfluous and tautological. We would also obtain verification in this case if the scale were graduated in any arbitrary non-uniform way, and the motion of the body were so manipulated by invisible strings as to satisfy the integrated equation.

In setting up the system in this way, with a long stationary scale and a falling clock, we have done more than is required for a verification of the differential equation. To do this it would be adequate to distribute along the path of the falling body a series of infinitesimal

SOME OF OUR PHYSICAL CONCEPTS—I

measuring sticks and, beside each infinitesimal stick, a clock adjusted to the correct rate but with an arbitrary zero setting. With this set-up, we could verify the differential equation, but not the integrated form. Verification is not now superfluous, but involves a check on whether the origins of the various infinitesimal measuring sticks have been so adjusted that $s = \int ds$, and the settings of the various clocks so adjusted that $t = \int dt$. (This discussion has obviously not been concerned with relativity considerations.)

It is possible to adopt an extreme view here and maintain that in no case is verification tautological or superfluous. This is the point of view of extreme sophistication which sees our whole conceptual method of handling the world, our recognition of objects with identity and our logical methods of handling them, as merely, at any epoch of time, a summary of past experience up to that date, with no guarantee of continuance and therefore with continual need of verification. I can see no logical method of refuting this position and I think it is essentially correct, but it is obvious that the physicist, or any one else for that matter, must operate on a different level to get anywhere.

We need not take as extreme a position as this to maintain that verification in this particular case is not superfluous, or in any other case where the mathematics has presented us with an integrated equation out of a differential equation. I think that many of the contemporaries of Newton or Leibnitz would have keenly felt the need of an experimental verification of the possibility of representing any actual situation by a differential equation. For the concepts for dealing with motion are not easily derived from the concepts with which static situations may be successfully handled, nor is the concept of a mathematical limit easy or its connection obvious with what we do in the laboratory. It must have been many years before our present mental serenity was acquired in the face of the operations of the calculus. It is perhaps possible even today to take the position that a differential equation is entirely an affair of the paper-and-pencil domain and therefore can never be subject to instrumental verification. I prefer to take what is perhaps a less rigorous point of view, and handle a differential equation, in such situations as that of the falling body, as something of which it makes sense to ask 'is it true?' It is, I think, in any event obvious that a 'law of nature,' or merely the results of specific observations, when embodied in differential form, of necessity contains a larger paper-and-pencil component than when expressed

in integrated form. There is no physical operation corresponding exactly to the mathematical operation of taking the limit. What happens in the physical situation is perhaps sufficiently suggested by our description above of a verification of the differential equation of a falling body in terms of closely spaced photographs with a high speed movie camera. In general, we make our instrumental readings as closely spaced as we can, and then make a logical jump to the paper-and-pencil domain by plotting these readings on paper on a greatly magnified scale and observing that we can draw a smooth curve through them. We regard the differential equation as verified if we can detect no consistent discrepancy between the smooth curve which we draw so as to satisfy the equation and the plotted experimental points. There is never any question of proceeding to a limit in the mathematical sense, and logically the hiatus between the paper-and-pencil operations of the calculus and our instrumental operations cannot be closed. We do not regard this hiatus as particularly serious, however, and think of a law of nature when formulated in differential form as something of which it has meaning to say that it may be directly verified.

Perhaps we have been making too much of the logical hiatus between the differential equation and the results of instrumental operations, because we have already seen that even when we express our results in integrated form there is of necessity a logical hiatus in that no instrumental result has the sharpness of a number in mathematics. It is merely that the failure of logical correspondence is perhaps somewhat aggravated in the case of the differential equation.

We now turn to a detailed analysis of some of our specific physical concepts. We shall particularly attempt to separate the instrumental from the paper-and-pencil or verbal component, and to see in just what way the paper-and-pencil or verbal component makes its eventual instrumental contact.

The first of the concepts to which we direct ourselves is the concept of the 'field.' The concept of the field is often presented as one of the cornerstones of modern physics, invented by the intuitional genius of Faraday, clothed in mathematical form by Maxwell, and crowned by Einstein in his general relativity theory. The great virtue of the field concept is usually stated to be that it absolves us from accepting that intellectual monstrosity, action at a distance. It is felt to be more acceptable to rational thought to conceive of the gravitational action of the sun on the earth, for example, as propagated through the

SOME OF OUR PHYSICAL CONCEPTS—I

intermediate space by the handing on of some sort of influence from one point to its proximate neighbour, than to think of the action overleaping the intervening distance and finding its target by some sort of teleological clairvoyance. That is, the intermediate space is pictured as the seat of some sort of physical action which propagates itself from point to point, eventually concentrating itself on its target, where it manifests itself in the form of a mechanical force. Now if this concept of something going on in apparently empty space is to be capable of eventual instrumental emergence and thus to have physical content, there must be some instrumental method of demonstrating that there is something going on. There seems to be only one conceivable method of demonstrating such a state of affairs, and that is to go to the point in question with a suitable instrument and observe that the instrument gives a reading. This we can in fact do. For our instrument we may use a small massive particle attached to a spring balance, and when we transport our instrument to the point in question the spring balance gives a reading, which by varying the orientation of the balance may be converted into a reading of a vector 'force' acting on the particle. Our point of view is therefore thus far justified. But have we really done what by implication we set out to do, and shown that the other point of view, of action at a distance, is inadmissible? Would not our instrument still give exactly the same reading as before if we had action at a distance? I think it is obvious that it would, and that as far as the instrumental criterion goes there is no distinction. This dilemma seems to be unconsciously recognised, and apparently resolved to general satisfaction, by defining the field at a point, not as the force acting on the actual exploring particle, but as something derived from the measured force, namely the limit of the ratio of the force to the exploring mass as the mass becomes smaller without limit. The limiting ratio obtained in this way is defined as the 'field' at the point, and is conceptualised as something characterising the point by itself, from which the effect of the test mass has disappeared because its magnitude has vanished. This is, it seems to me, plainly not a correct way of conceptualising the result of going to the limit, for the physical effect of the test mass persists throughout every stage of the limiting process, the force on the mass finally vanishing together with the mass. The effect of the mass has in nowise disappeared in the limit, since we are taking a ratio of two quantities, both approaching zero. The device of taking the limit thus appears as unsuccessful in accomplishing the desired purpose.

There is, of course, a legitimate reason for using the limiting process in defining the field at a *point*; the limit may be necessary to smooth out the effect of space variations of the field if the field varies rapidly from point to point and the test mass extends over a region wide enough so that the variation is perceptible. Also the limiting process is necessary to eliminate the reaction of the exploring instrument back on the distribution which gives rise to the field; this remark is particularly applicable when the field is an electric field arising from a distribution of charge on conductors.

Furthermore, a serious logical difficulty appears when the field concept is carried through in mathematical detail. For it appears that the *mechanism* by which the field exerts a force on the charge is by way of the perturbations which the test charge automatically introduces in the field. These perturbations are the essence of the matter because they are proportional to the test charge and account for the entire force. Thus even in the paper-and-pencil domain of mathematical manipulation we can find no operational meaning in terms of mental experiments of a field undisturbed by the instrument of measurement.

It seems to me that there is no way by which the desired distinction between action at a distance and action by a field can be given instrumental significance. For always the instrument by which we would establish the existence of the field is subject to the suspect action at a distance. The situation which has thus presented itself is part of a more general situation, for always, from the point of view of operations, it is fruitless and meaningless to attempt to establish the existence of anything independent of the means by which its existence is established or verified. The two together, object and means of observation or measurement, form an indissoluble union; either without the other is meaningless, at least in the instrumental domain. Yet from the point of view of common sense it is incontestable that we do conceptualise the objects of daily life as having an existence independent of the instrument or method of observation. In the commonsense domain the meaning of this is clear enough; it is that we are able to adapt ourselves to everyday objects, that is, to draw our programmes of action involving the objects, without taking explicit account in those programmes of the methods by which we acquired our knowledge. Since the method of observation is irrelevant for our usual purposes, it is economy of thought to forget it, and we think of objects as having an existence independent of observation.

SOME OF OUR PHYSICAL CONCEPTS—I

This commonsense view comes to colour all our thinking and we carry it over uncritically into situations not sufficiently like those of common sense to justify the extension.

Returning now to our analysis of the field concept, ostensibly the device of making our exploring charge vanishingly small in the limit provides the means of achieving independence of the measuring process, but actually, as we have seen, it does not, the presence of the charge itself, even when vanishingly small, always being an essential element in the situation. The situation might conceivably be improved if there were *two* or more independent instrumental means of verifying the existence of the field, such for example as the electrostatic double refraction which accompanies an electrostatic field in a transparent body. But apparently there are not two such independent instrumental approaches in empty space, but the only instrumental meaning of the field is in terms of the force on test bodies. Such other effects as double refraction occur only in the presence of matter, which automatically provides the possibility of action by contiguous contact. But even if there were two or more independent means of demonstrating the existence of a field, it is not at all obvious that we could then show that action at a distance is not an alternative method of description.

Although we apparently cannot have two independent instrumental methods of directly verifying the existence of the field or showing the impossibility of action at a distance, there may be other sorts of phenomena and other sorts of instrumentation that may, in a wider setting, make the field concept under certain circumstances so much more convenient to handle than that of action at a distance that we may use it exclusively. Such independent phenomena do exist in the case of the electromagnetic field, so that the whole picture is different for electrical than for gravitational phenomena. We will return to this question in connection with the concept of energy and its localisation.¹

However repulsive the concept of action at a distance may be to our mechanical intuition, it does not violate the broadest demand which we can put on the sort of account we must give of the external world if we are eventually to reduce the external world to understandability. It seems to me that the broadest basis on which we can hope for an eventual understanding is *invariable correlation between the*

¹ The lecture in which this is discussed will be published in the next number of this *Journal*.

results of instrumental operations. Given invariable correlation, we can find how to predict, and prediction is perhaps the most searching criterion of understanding. We can have such invariable correlation either in terms of a field or of action at a distance.

The fundamental situation with which we are confronted here and the demands which we automatically exact of our treatment of the situation come perilously close to involving an inner contradiction. We want to talk about a field in otherwise empty space, and we demand that we have some instrumental indication of the existence of the field in spite of, or ignoring, the fact that the space can no longer be empty when we introduce the instrument of detection. The failure of the space to be empty in the presence of the detecting instrument is not an irrelevant effect that can be neglected, for how shall we rule out the possibility that the detecting instrument has introduced something with it, indissolubly tethered to it, like the physical lines of force of J. J. Thomson? To deduce an objectively existing field from instrumental measurements of a force which is proportional to the test charge involves the same sort of situation as if in the ordinary realm of macroscopic phenomena the mass of a body were proportional to the intensity of the light with which we observe it. The concept of independently existing objects certainly would not have formed itself under such conditions. So here, it is only by an uncritical analogy that we form the concept of a field independently existing in its own right. Apparently all that nature will grant us here is the instrumental reaction when we station our instrument at any point in formerly unoccupied space; she will not grant us the further right to analyse our experience into a field as distinguished from action at a distance. Instrumentally the distinction between field and action at a distance appears to be meaningless. We must accordingly recognise that the distinction which physicists actually do make between these two concepts is verbal, and the corresponding operations are verbal operations. There is no doubt that verbalisations about the distinction between the field and the action at a distance points of view have played an important part in the activity of physicists, as may be seen by consulting the literature. Doubtless with our present mathematical machinery it is much more convenient, particularly when dealing with electromagnetic phenomena, to prefer the field to the action at a distance point of view, but one may feel a certain scepticism as to whether such distinctions really play an essential role.

We have been talking about 'empty' space in connection with the

SOME OF OUR PHYSICAL CONCEPTS—I

field concept. In the concept of 'empty' space we are again coming perilously close to an inner contradiction. For how shall we establish that a purportedly empty space is really empty without going there with an instrument to prove it, and when we have introduced the instrument the space is no longer empty. In the realm of macroscopic experience with ordinary matter we can deal with similar situations by setting up a theory which takes account of the effect of the instrument so that we can correct for the perturbations introduced by the instrument, as when we correct for the effect of the size of the bulb of a thermometer on temperature readings. The theory by which we correct for the perturbations is based on the variation in the behaviour of the instrument when it is applied to a range of situations. But in the limit when the matter vanishes but the instrument does not, such procedure fails us, while the intellectual compulsion remains to give some instrumental meaning to the purported emptiness of space. The simplest way of meeting this compulsion is simply to say that the space is empty if *no* instrument gives any reading when introduced into it. But are there such empty spaces? Up to a short time ago it would have been accepted as intuitively obvious that it was permissible to use the concept of empty space in our speculating with no danger of ever encountering inner contradiction. In fact, the concept of empty space seemed almost a necessity of thought. A sufficiently critical insight might have seen, however, that the question of the 'real existence' of empty space was in some way bound up with the behaviour of instruments, and so became a question of experiment rather than of *a priori* logic. And now within the last few years we have the quantum mechanical concept of a fluctuating zero point electrostatic field in otherwise empty space. If the theory is correct, it means that it will be found as a matter of experiment that it never occurs that there are places where all physical instruments give no readings, so that 'empty' space corresponds as little to the physical actuality as do the simultaneous position and momentum prohibited by the Heisenberg principle of interdetermination. In this denial of the legitimacy of the concept of empty space it seems to me that we have as dramatic a demonstration as can be imagined of the impossibility of divorcing our concepts from the operations by which they are generated and of the impossibility of speaking of things existing of themselves in their own right.

There is another class of physical phenomena in which we form our concepts by ignoring part of the operational background, namely the

phenomena of the propagation of light or radiation in general. There have been a number of different physical pictures of the nature of light, all of which have had the feature in common that light is to be regarded as in some way a 'thing travelling.' This statement obviously is applicable to the old corpuscular theory of light; it is true for the electromagnetic theory where the 'thing' is a phase in the electromagnetic field, which may be followed in thought as it moves, as in Einstein's special relativity; or finally we have the 'thing travelling' picture in the photons of quantum mechanics. What instrumental method shall we adopt to show that when we have light we also have a thing travelling? In the realm of ordinary optical phenomena we are here confronted with the same sort of self-defeating requirement that we have already met in trying to distinguish instrumentally between the field and action at a distance or in trying to give instrumental meaning to empty space. For the only method of detecting the presence of light is to put a screen in the path of the radiation and observe that the screen is illuminated. We never experience light as such but only things lighted; operationally, light means things lighted and not a thing travelling. Any meaning that can be given to the thing travelling concept can be only an indirect meaning, involving other sorts of operation than those contained in the definition. What other sorts of operation would be acceptable? In the case of ordinary travelling things, such as a baseball in flight from pitcher to catcher, there are other phenomena besides those at the delivering and receiving ends, such as being able to see the ball in flight or detecting the wind as it passes. But we cannot see a photon in flight nor does it create a wind. All that we have is certain geometrical relations between things lighted which are the same as if there were rectilinear propagation (neglecting here diffraction and similar effects). There is, however, no instrumental method of proving that 'action at a distance' does not also follow rectilinear rules. I believe that the situation is not altered by considering the *velocity* of light, for there appears to be no reason in principle why action at a distance should not take place at more distant points at later instants of time, thus involving the instrumental attributes of a velocity of propagation. Until new sorts of experimental fact are discovered it seems to me that the concept of light as a thing travelling remains a predominantly paper-and-pencil concept, mostly verbal in character. This verbal concept is of undoubted value because it enables us to make our mental experiments and conduct our paper-and-pencil operations in a congenial fashion

SOME OF OUR PHYSICAL CONCEPTS—I

closely analogous to the way in which we treat ordinary material macroscopic objects. But I doubt whether there is any logical necessity here, and I believe that all the results could also be obtained with a theoretical apparatus that pictured light as some sort of delayed action at a distance.

The proviso above 'until new sorts of experimental fact are discovered' is an essential part of the situation, and indicates that purely verbal concepts may nevertheless have instrumental implications and therefore not be indifferently equivalent to each other if we consider them in the context of the programmes of experiment which they suggest. The application here is immediate. For if light 'really' consists of photons travelling and these photons have physical 'existence' then we would anticipate some sort of interference between the photons of two beams of light crossing each other in otherwise empty space. Theory should be in a position to calculate the amount of such interference to be expected. The calculations have been made and it proves that the amount of interference to be expected is too small to be detectable with light intensities at present attainable. If and when such interference effects can be instrumentally detected the concept of light as a thing travelling will have been so far justified.

There is another way of dealing with light which is closely related to the thing-travelling point of view, namely to speak of light as going *through* space. Again we meet a self-defeating situation, for any instrument set to detect the passage of light by its very presence prevents the passage. Instrumentally, light only departs and arrives. The emphasis in speaking of light going through space is different from the emphasis in speaking of a thing travelling, for now we direct attention to the space, and presently are speaking of a propagation of light through space, and of space having the *property* of propagating light with a definite velocity. At first glance it is as paradoxical to assign properties to empty space as it was a few pages back to discover that no instrumental meaning can be given to empty space. The two points of view now appear merely as two aspects of a wider vision, namely that since what we shall find by instrumental manipulation is wholly a matter for experiment to decide, we cannot have unlicensed freedom in our paper-and-pencil concepts if we also demand that our paper-and-pencil concepts eventually emerge into the instrumental world.

Formally, and taken in isolation, the point of view of light as a thing travelling or propagated cannot be justified in preference to the

point of view of light as a thing departing and arriving with retardation in time, so that whichever point of view is adopted it must be described as a convention. It may well be, however, that one of two alternative points of view is so much more congenial to the commonsense way of looking at things, the commonsense point of view itself being recognised as at bottom a construction, that we shall adopt it in preference to the other. It may well happen that in an early stage of knowledge there was little to choose between two points of view but with the discovery of new experimental facts the convenience and simplicity of one point of view comes to be so overwhelming as to result in the unanimous discarding of the other. This is what happened with our concepts of atoms. Fifty years ago there was a serious school led by Ostwald, that maintained that the concept of atoms was superfluous, since all it expressed was the fact of constant combining weights in chemistry. But with all our new experimental discoveries, the implications of Brownian motion and tracks in a Wilson condensation chamber, to mention only two, the convenience of the atomic picture has become so overwhelming that we have discarded the alternative point of view completely, and speak of the situation in different terms, as when it is often stated that all these facts have proved the *reality* of atoms. I suspect, however, that the logical situation has been in no way altered by all the new experimental discoveries, and that the old Ostwald point of view could be carried through if we were willing to pay the price in complexity. Incidentally, this example suggests what the operational meaning of *reality* is. It seems to me that very much the same sort of thing that happened with regard to our concept of atoms is now happening with regard to our concepts of light and field and empty space and action at a distance. With the discovery of new facts, such as the creation of electron pairs in apparently empty space, the one point of view becomes so much more convenient that we forget the possibility of the other and discard it. But I think we should not forget what we are doing here and remember that logically we have nothing unique but are adopting a particular convention because of the naturalness with which it enters our commonsense scheme of thought.

P. W. BRIDGMAN

CAUSATION AND EXPLANATION IN THEORETICAL BIOLOGY

I *Introduction*

WHEN it is seen that causal explanation can never be the ultimate aim of scientific study, it should be possible to remove some of the misunderstandings which have produced 'schools' in theoretical biology. It is perhaps a mistake to suppose that causal study belongs merely to the 'infancy of a science,' but causal investigations are certainly always stages in our progress towards the discovery of ultimate regularities, and when these are reached it is a mistake to seek *further* causal explanations.

Thus, if rectilinear motion be the motion which all bodies tend to assume, gravitation appears as a cause explaining deviations from this general law. But unless gravitation itself can be shown to be a deviation from some more general rule, it is superfluous to require a cause for gravitation. Fourier, then, was mistaken in saying 'Les causes primordiales ne nous sont pas connues,' for causal phenomena cannot be primordial, if they are causal. Or perhaps it would be more correct to say that the last-known regularities may *be* causes (as in the case of gravitation) but cannot *have* causes. In this case, 'les causes primordiales' would be the last regularities we are able to discover.

We may, of course, be mistaken in our notion of what constitutes an ultimate regularity: Aristotle appears to have thought that rest is an ultimate regularity, and that uniform motion is a deviation from this, which requires a cause.

Science, then, is not ultimately aetiological, but nomological.¹ This distinction appears to be of importance in three related questions in theoretical biology:

- (a) What place are we to assign to teleology in biological theory?
- (b) To what extent is it reasonable to hope that the molar phenomena of organisms will ultimately be explained as resultants

¹ I have avoided the word 'aetiology' elsewhere in this paper, because of its special meaning in medical science, and because it seems to me that the word *αἰτιον*, at least as it was used by Aristotle, means any sort of explanation, whether causal or nomological.

of molecular? That is, has the reductive science of organisms logical limits?

(c) What importance is to be attached to mnemonic phenomena?

It will be seen that although the first two questions are related, they do not present alternatives, for the impossibility of molecular explanation would not necessarily mean the introduction of teleological; nor are the two types of explanation mutually exclusive, for molecular explanation does not exclude teleology. In Aristotle's example, we cannot deny that the lantern can be explained by the properties of its parts, nor does this explanation make it untrue to say that the man who carries, or the man who makes, the lantern does so for use. Again, Aristotle's explanation of the ink of the cuttle-fish, which Ross finds ludicrous, does not strike the Darwinian as at all odd.¹

The second and third questions are related, for mnemonic phenomena are among the facts which set limits to the possibility of reductive explanation of organisms. Also the first and third are related, for it is the fact that organisms can 'make use of' memories which gives memory a special place among mnemonic phenomena.

2 Teleological Explanation

We can best examine teleological explanation by starting from an example in which the element of conscious purpose is certainly present, the better to distinguish those cases in which we cannot assume this to be the case.

If we ask, 'Why did you blow the fire?' we are asking for a purposive explanation, the question meaning, 'Of what intention were you aware when you blew the fire?' We do not ask about unconscious motives, which might in fact be the more interesting and important, but which cannot be elicited by direct questioning. Direct questioning is, however, the only way of eliciting conscious intention.

It is clear that the answer to such a question need have very little to do with the fire. If a child asks the question, we may answer,

¹ Aristotle supposed that the ink of the cuttle-fish was at first a mere excretion, but that it became useful to the fish in hiding it from its enemies: he went so far as to point out that it is natural for many animals to void their excrement when startled, even when this can be of no use in concealment. Ross says of this sort of explanation that the utilisation for an end is 'a sort of afterthought on nature's part' (W. D. Ross, *Aristotle's Prior and Posterior Analytics*, revised text and commentary, Oxford, 1949, p. 645; Aristotle, *de Partibus Animalium*, 679 a, 25-30.)

THEORETICAL BIOLOGY

'To make it burn better,' but if the question is put by an adult it is probably a purely psychological one. The answer in this case might be, 'A nervous habit,' or 'To hide my embarrassment,' or 'To express my irritation.'

On the other hand, it is possible to enquire, as a child might, why blowing the fire makes the fire burn better? Now this might be a causal question of the familiar type ('efficient causation'); we might consider the act of blowing as a physical cause and the acceleration of combustion as the effect. But it is possible also to be guided by a somewhat different principle. Blowing a fire is only a special case of providing a draught, and this is a necessary condition of all ordinary combustion. If we consider a fire, or a furnace, as a 'relatively isolated system,' it is possible to ask what part is played by a draught in such a system. This question depends for its legitimation on the recognition that the draught is an integral part of the system; it is not, as the blowing of a domestic fire is, an occasional interference with the system from without, but a necessary and constant part of the system. With respect to the system, therefore, it is not a cause, and the question we ask about it is not strictly a question about causation though a nearly related question can, as we have seen, be treated as one of causation.¹

In the case of the fire, this distinction may appear trivial. Neither its isolation nor its internal structure is of a very high order. When, however, we encounter more elaborate systems of the same kind such as furnaces, it becomes clearer that the question 'What is that part for?' is not necessarily answered in the same way as the question 'How does that part work?'—but both are legitimate questions, and neither has any direct reference to purpose. When we ask the engineer, 'What is that for?' we are not really asking the same thing as 'Why did you put that there?'—the question *might* be put in the second form, but it would sound odd. Of course, when we are questioning an engineer, the questions 'What purpose had you in view?' and 'What is the function of that?' are very closely related, but we have seen in the discussion of blowing the fire that they are not identical, and it seems to me that we may legitimately enquire as to the part played in the economy of the system, without any reference to what could properly be called 'purpose.'

¹ When I say that a draught is a 'necessary' part of the system, I mean, of course, necessary by definition. A 'fire' without a draught (an active thermite bomb, for example) would be quite a different kind of system, with different laws *qua* system.

When we turn to the study of living things, we find that the distinction between a causal and a functional explanation is recognised implicitly in the way we usually speak of the parts of animals. For example, to ask, 'What is the function of the heart?' is a normal question, and 'What are the functions of blood?' might appear in an elementary examination paper for students of biology. But 'What is the effect of the heart?' or 'What are the effects of blood?' are questions which have a very odd sound, and which certainly would never appear in an examination paper.

It may be helpful at this stage to review what is common knowledge about the heart, so that we may be enabled to put the teleological explanation in its proper perspective.

(a) *Causal Knowledge of the Heart.* We have become so accustomed to the idea that the blood is caused to circulate by the action of the heart, that the knowledge has ceased to arouse our interest, but we must not forget that this discovery was one of the turning points in physiological science. Without this knowledge, neither the mechanism of the heart nor its function in the economy of the organism would be comprehensible. This is perhaps an illustration of a general feature of scientific progress, especially in the biological sciences: that causal study must precede nomology; unless nomology is to remain at the stage of classification which is usually supposed not to have explanatory value.

Most of what is most properly called the physiology of the heart consists of knowledge of causes and effects within the organism. The knowledge of the nervous control of the heart and of the effects of hormones or other chemical substances on the cardiac muscles is strictly causal, because there undoubtedly is a basic heart-rate which is accelerated, depressed, or rendered irregular by these agencies, which can consequently be considered as disturbing or interfering processes and thus fulfil the definition of causes. It is of course quite clear that pathology, by its nature, is a causal science, since it supposes a standard 'condition of health' which is disturbed by the pathogenic agents. This is illustrated graphically in electrocardiography, which is a method of recording disturbances (effects) with the intention of identifying disturbing agencies (causes). One interesting piece of causal knowledge about the heart still remains to be discovered since, so far as I am aware, we do not seem to know what first makes the heart beat in the young embryo.

(b) *Nomology of the Heart.* I am not sure whether the mechanical

THEORETICAL BIOLOGY

explanation of the heart (the mode of functioning of the valves, the relative thicknesses of the ventricular walls, etc.) can be considered in any way causal. I can appreciate that once the mechanism is set going at a regular speed, variations in that basic speed are effects which require causes, but I cannot see that the exemplification of general mechanical laws has anything causal about it. The particular movement of a particular lever, which we know to have been long at rest is a causal event ; but levers as such are indifferent to rest or motion, and as soon as we leave the particular case and generalise our mechanics it seems to me that we leave causal studies behind. However, I am not sure where we should draw our distinctions in this case. Personally, I do not regard mechanical explanation as causal. A second difficulty arises over the microphysical explanation of the action of heart-muscle. When we can say exactly what molecular process it is which finds its molar expression in muscular contraction of a rhythmic kind, we shall be able to give a very satisfactory explanation of the basic action of the heart ; but I cannot myself see any reason for considering that this would be a causal explanation. Aristotle would say that we were assigning τὸ αἰτιον of the heart's action but it appears that he did not distinguish explicitly between causal and nomological explanation.

Finally, I should deny that the teleological explanation of the heart is a causal one. It appears to me that teleological explanations contain an element of deduction which is entirely absent from causal. The only deduction we can make from effect to cause seems to be of this nature :

If a process has been regular (such, in the ideal case, that we can express it by a relatively simple formula) between times t_0 and t_1 , and if after t_1 it no longer can be expressed by the same formula, then, either we are mistaken in supposing we have in any sense the same process after t_1 as before, or, if we are justified in saying 'it is the same process, but modified,' then we must introduce some new factor into our formula to express the process as it is after t_1 . If, then, we are ever justified in saying that we have an instance of the modification of a process (as for example, the modification of the speed of the heart), it follows that this modification must remain unexplained, or must be explained by the introduction of some new factor ; and this is the justification for arguing from effect to cause. But this tells us nothing of the nature of the cause in any physical terms. If the heart beats faster, all we know of the cause is that when it occurs the heart beats faster.

We are justified then, in saying that an effect cannot be explained without a cause, but this seems to be as far as deduction will take us in general. The only case in which we can argue from effect to cause is that in which we consider we have an entirely specific effect of which the cause is known : ' if this is tuberculosis, it is caused by the tubercular bacillus.' This argument either depends on the validity of inductions in general, or it is tautological (' if this is not caused by the tubercular bacillus, we shall not call it tuberculosis ').

Teleological deduction seems to be different in kind from this, and is certainly different in degree. What we mean by a teleological explanation of the heart is that, given an animal of a certain size, we can deduce that it must have a heart ; that is, some apparatus for pumping blood. This argument depends partly on induction, since it supposes a certain speed of metabolism and certain limited means of maintaining that speed, but it depends also on a mathematical proportion ; the fact that the *ratio* of volume to surface in a compact body varies directly as the radius.

It seems to me that this explanation of the heart which shows the place of the part in the whole system, which says in effect, ' Given such a system, there must be such a part,' is as satisfying as the mechanical explanation of the manner in which the part performs its necessary function.

This example of teleological explanation may give rise to the supposition that teleological explanation is merely mechanical explanation with the terms transposed. Is not the difference perhaps as trivial as that between $y = 2x$ and $x = \frac{1}{2}y$? Let us consider the following four questions :

- (i) What load will a given bridge (i.e. of known construction and materials) take ?
- (ii) What sort of bridge will be required to take a given load ?
- (iii) How does the jaw-apparatus of a mammal triturate food ?
- (iv) What kind of apparatus is necessary for the trituration of food ?

It will be seen that questions (i) and (iii) will be answered by a mechanical explanation of particular structures, whereas (ii) and (iv), though in a sense equally mechanical, do not specify the particular mechanism which will satisfy the conditions. Even if the materials are fixed, several designs may be proposed for the bridge, and food may be triturated by mechanisms as different as the jaws of a mammal

THEORETICAL BIOLOGY

and the crop of a bird. But there is a further distinction to be made. In questions (i) and (ii) the final reference is to a load which is not part of the bridge, and the questions are complete in themselves. But questions (iii) and (iv) refer to something (food) which is about to become a part of the system. Question (iv), therefore, which is cast in a teleological form, suggests further teleological development.

We may enquire concerning the meaning of trituration of food.¹ It is a perfectly sensible and scientific question to ask, what part is played by trituration of food in the economy of the organism? We may continue our explanation of the trituration of food by saying that triturated food is more rapidly and efficiently digested. An animal with a rapid and efficient digestive system will get more energy out of each meal, and can take more meals in a given number of hours. We can then terminate our analysis with the notion of the more efficient machine, or with that of the animal having the greater survival value.

This sort of analysis may be quite unnecessary for the physiologist, but if the general biologist finds it useful, I cannot see on what grounds it can be denied to him. It is the only explanation which allows us to connect the jaws with the rest of the alimentary system in a comprehensible manner; and if it is no longer useful in the advance of the science, that is because, like the circulation of the blood, we now take it for granted.

The difference between the mechanical and the teleological explanation becomes particularly clear, and the teleological explanation ceases to be academic, when a part is present in one group of animals, but absent in a related group. It is clear that this cannot be explained by any knowledge of the part as such, but requires a consideration of the place of the part in the whole, and of the way in which the different structure or different circumstances of the two classes is correlated with the presence or absence of the part. Thus a part A may be absent because a part B is absent; or the part A may be absent because the circumstances of the two groups are different.²

If the distinction between causal and nomological study still

¹ The same situation would, of course, arise if our questions about the bridge had concerned, not the load which the whole structure is adapted to carry, but some part of the structure.

² A particularly instructive example of this sort of study is to be found in E. Baldwin's elegant *Introduction to Comparative Biochemistry* (Cambridge, 1940).

appears somewhat academic, it may be helpful to consider whether we do not see an effect of neglecting this distinction in the faults of emphasis due to an exclusively pathological approach to some problem (pathology, as has been pointed out, is a predominantly causal science). Perhaps in the early history of endocrinology we may see the disturbing effects of attaching too much importance to the pathological activity of the glands, so that their function came to be regarded as episodic, and the glands themselves as acting upon the system rather from without, than as forming themselves a constant part of the system. Certainly, most academic psychologists would consider that the psycho-analysts had been misled by their pathological approach. In spite of their efforts to the contrary, they often give the impression that instincts act from without, and causally, upon the system. The difficulty is, I think, that a causal event is a compound AB, and that if we are interested in B, AB may tell us a great deal about it, but there is always some danger that we may come to think of B as being identical with AB; this danger will, of course be the more real if neither A nor B is a sensible entity, or, as we might say, a construct of the first order.

3 *The Limits of Reductive Explanation*

I *Reductive Explanations*. It might be thought that, if we abstract from mnemonic phenomena, and consider the organism as a relatively isolated system, we should have good hopes of being able to explain its activities, ultimately, in the language of physical chemistry.¹

Dr Woodger has discussed this subject at some length,² and with his usual lucidity, so that not a great deal remains to be said about it. Dr Woodger seems to be of opinion that the mechanical explanation of living things was introduced as a result of the failure of the attempt to produce a tolerable physico-chemical explanation. I do not think this is true historically, since there seems no doubt that the mechanical theory is the more ancient of the two, but this is a point of no very great importance. What is important is to observe that both methods are necessary, and that, though both are equally 'physical,' they do not reduce to the same thing.

The contrast between the two is very clearly seen in the two parts

¹ 'The intricate network of psychological intercorrelations, when expressed in the physico-chemical language, will be unintelligible to the layman' (Carroll C. Pratt, *The Logic of Modern Psychology*, New York, 1948).

² J. H. Woodger, *Biological Principles*, London, 1929

THEORETICAL BIOLOGY

of D'Arcy Thompson's classical book.¹ In the earlier chapters we see how it is possible to explain much of the structure of the smallest organisms, and many small parts of the larger, in terms of micro-physical laws. But when D'Arcy Thompson comes to the explanation of the skeletons of the higher animals, the same method cannot be applied.

The mechanical laws which govern the skeleton of an ox are those which also govern the construction of a bridge, and the explanation of the larger skeleton is as scientific and as satisfying as that of the skeletons of the micro-organisms such as radiolaria or foraminifera. But between the molecular and the molar explanations two points of difference are to be observed :

(a) One is easily convinced that the smaller structures, whether skeletal structures or cleavage-patterns of eggs, are as they are because of the properties of the materials of which they are composed. In the case of the cleavage-patterns it seems clear that the possible variants can be deduced from a knowledge of the properties of surface-films ; that all the possible variants in fact occur, and that they occur with a frequency which is much what might have been calculated. Further, these structure-patterns can be found in groups of soap-bubbles, which are not only non-living systems, but also natural systems (even if they happen to be produced artificially for the purpose of experiment).² But in the case of the skeletons of the larger animals there seems no possibility of maintaining that their mechanical structure arises from the properties of their materials. In this case, the properties of the materials only impose very general limits on the structures which can be produced.³ No one can doubt the justice of D'Arcy Thompson's comparisons of skeletons with bridges, but the materials are manifestly extremely diverse. Nor could anyone maintain, that all the possible variants are known if the laws of the materials are known.

(b) In the case of the radiolarian skeleton, to take one example from the smaller animals, we are entirely at a loss to suggest any mechanical or any functional explanation of the structure. If some skeleton be necessary to the animal, which seems improbable, no plausible suggestion can be made to account for the regular and elaborate conformations

¹ Sir D'Arcy W. Thompson, *On Growth and Form*, 2nd ed., Cambridge, 1942

² This is of importance for the evolutionist : since we may suppose natural structures to have been in existence before the appearance of animals, whereas artificial structures are themselves products of the animals.

³ Cf. C. D. Broad, *Mind and its Place in Nature*, London, 1947, p. 108

which we in fact find. In the case of the skeleton of the higher animal, however, the mechanical and functional explanations are inescapable, or at least can only be escaped by ruling them out of court as 'unscientific.' In fact, if we do not like them, we have to shut our eyes to them by an effort. Even if, therefore, we were to discover that the materials of a horse or an ox are such as to give rise to mechanical models in exactly the way in which soap-bubbles give rise to cleavage-patterns, we should still be able to give another explanation of the structures in perfectly general mechanical terms, in no way dependent upon the nature of the particular materials involved.

II Logical Status of Reductive Explanation. It can, I think, be shown that reductive explanation, at least within certain limits, is logically impossible. Let us consider the explanation of the muscles of the lower jaw of a mammal. As far as I am aware we cannot say yet exactly what is the underlying process in the contraction of skeletal muscle, but there is no reason why this should not soon be discovered. Let us suppose that the contraction is an expression of a change of forces between the fluid sarcoplasm and the fine tubes in which it lies; a phenomenon of 'capillarity.' Now if this is an explanation of the contraction of skeletal muscles in general, it is clear that it must fail to explain any particularity which is possessed by any muscle over and above being a skeletal muscle.

If we now say that the particular muscle we are considering is part of a mechanical structure, and is explained as the pull applied to a lever of the third order, we clearly add something which cannot be deduced from the general theory of muscular activity; for there is nothing in the theory of muscular contraction to specify its mode of application, and in fact it is applied to other levers in other ways. Moreover, just as muscles applying force to levers of the third order are a sub-class of muscles in general, so the temporal muscles are a sub-class of this sub-class; for there are other such levers and muscles in the body (for example, the biceps as applied to lifting the fore-arm against gravity).

If, therefore, we wish to continue our process of specification, we must explain the temporal muscle as such. Neither of the preceding explanations is sufficiently specific, and we must refer to the particular part and its particular function (mastication etc.). We can now see how this type of 'teleological' explanation remains necessarily

THEORETICAL BIOLOGY

irreducible, and why we must reconcile ourselves to it or have no explanation of the *particular* part.

The above analysis seems to show that the levels of explanation are irreducible because they have the relation of genus to species to one another. We might, of course, achieve a reductive explanation if we could show that each and every muscle was characterised, over and above the general physical chemistry of muscles, by its own specific chemical composition ; but this would remain an explanation of a very low order, indeed hardly more than a fact without explanatory value, unless we could show what properties of the specifying chemicals were responsible for the specific peculiarities of the particular muscles. An explanation, after all, must generalise, and this is done by all of the three explanations we have mentioned above : there are capillary phenomena elsewhere than in muscles, levers elsewhere than in the body, and, if masticatory structures are found only in animals, they are at least widely distributed and of various patterns.¹

Note : Kant considers that the reductive explanation *is* the mechanical explanation, and that any explanation which is not reductive is teleological. But this leaves us with a third type of explanation unaccounted for. The teleological explanation of a bird's wing is simply that, with respect to the bird, the wing is an organ of locomotion.²

The most satisfactory explanation is that which connects the form of the wing with the act of flying, and this, I suppose is what most people would call the 'mechanical explanation.' The reductive explanation (Kant's 'mechanical explanation') would require that the form of the wing should be deduced from the properties of its constituent materials, and this would suppose that the materials of the wing are other than those of the leg, which does not seem to be the case. The relevant passage is to be found in the *Kritik der Urtheilskraft*, section 77 :

¹ The reader may wonder why I have made no mention of the Aristotelian distinction between material and formal causation. This omission is partly motivated by the desire to discuss the subject in a modern idiom, but is due more especially to my belief that we are not dealing here with causal explanation at all. It seems to me that the prejudice against teleological explanation is partly due to a supposition that it is a causal explanation.

² This appears trivial in the case given, but anyone who has considered the various structures of the more complex Ciliates will certainly not consider it without interest to be told whether a particular 'bristle' is or is not an organ of locomotion.

Wenn wir nun ein Ganzes der Materie, seiner Form nach, als ein Produkt der Theile und ihrer Kräfte und Vermögen, sich von selbst zu verbinden (andere Materien, die diese einander zuführen, hinzugedacht), betrachten, so stellen wir uns eine mechanische Erzeugungsart desselben vor.

4 *Mnemic Phenomena and the Relatively Isolated System*

The novelty of Russell's theory of Mnemic Causation has perhaps somewhat obscured the importance of mnemic phenomena in general, of which Russell himself says, 'I think it is this characteristic, more than any other, that distinguishes sciences dealing with living organisms from physics.'¹ This does not mean that mnemic phenomena are found exclusively in living systems, and Russell quotes the magnetisation of steel as an example of this kind of phenomenon in a non-living system. Mill discussed this mode of causation, though he does not use the expression 'mnemic':

There is a case of causation which calls for separate notice, as it possesses a peculiar feature. . . . It often happens that the effect, or one of the effects, of a cause, is not to produce itself a certain phenomenon, but to fit something else for producing it. . . . Physiological agencies often have for the chief part of their operation to predispose the constitution to some mode of action. To take a simpler instance . . . : putting a coat of white paint on a wall does not merely produce in those who see it done, the sensation of white; it confers on the wall the permanent property of giving that kind of sensation. . . .

We must . . . include among the effects of causes, the capacities given to objects of being causes of other objects. This capacity is not a real thing existing in the objects; it is but a name for our conviction that they will act in a particular manner when certain new circumstances arise. We may invest the assurance of future events with a fictitious objective existence, by calling it a *state* of the object. But unless the state consists . . . in a collocation of particles, it expresses no present fact; it is but the contingent future fact brought back under another name.²

It would be most interesting to compare this passage in detail with Russell's treatment of the same subject, but we must content ourselves with a few observations. Both authors seem undecided whether or not we need a 'trace,' but both appear to think that a trace, if it existed,

¹ Bertrand Russell, *Analysis of Mind*, 5th ed., London, 1949, p. 83

² J. S. Mill, *A System of Logic, ratiocinative and inductive*, London, 1875, Vol. 1, pp. 388 ff.

THEORETICAL BIOLOGY

would consist of 'a collocation of particles' (viz. in the brain). Mill, it seems to me, quite unnecessarily brings in a reference to the future, as though sharing the notion that it is the business of science to produce prophecies. I do not see why we should not be contented with a behaviouristic criterion in the present. If the present behaviour of the animal satisfies us that the animal behaves otherwise than it would have done had it not been exposed to certain circumstances in the past (in plain English : if the animal's behaviour seems to be influenced by a memory) we can say that it retains a trace of the past experience, meaning no more than that its behaviour *does* appear to have been modified by a past experience. It seems to me we should be content with this sort of criterion in the case of any other system, and there is certainly no need to introduce a hypothetical future. No doubt, if we could find some sort of correlate, some 'isomorph' in the brain, we should be delighted, but it seems quite gratuitous to demand this before we consider ourselves to be justified in speaking of a trace.

This question is, however, as Russell is the first to admit, of no very great importance. What is of importance for our present discussion is that, whether we find or hope to find a trace in the nervous system or are content with a behaviouristic criterion, we are here concerned with a modification of a system which we treat as 'the same system' before, during and after the modification. The system is, therefore, in exactly the same case as a particle which is said to suffer a temporary or permanent deviation of the direction of its path. We have something which is taken as a regular process, and that process appears to suffer a modification through fusion or collision with another, and this is precisely what we mean by a causal event. Mnemic events, then, with respect to the organism considered as a relatively isolated system having its own nomology, are properly speaking causal events.

The distinctions which I have here adumbrated may be illustrated by reference to soap-bubbles. It is clear that soap-bubbles are relatively isolated systems whose conformations are explicable in terms of certain relatively simple physical laws of pressure and tension. The laws are not peculiar to the system (the bubble) but they are in a manner proper to it ; in the sense that the form of a bubble, or of a number of contiguous bubbles, can be explained without reference to anything outside the system, provided we are acquainted with these general laws. The system has in this sense its own nomology. This is true,

however, only in the most extreme case of the relative isolation of the system ; say, when the bubbles are floating in still air. As soon as the system of bubbles comes in contact with a rigid surface, the internal laws no longer suffice to account for its conformation. The contact, accidental and historical to the system as such, introduces a causal element into our study. The form of a bubble floating freely has no cause, it is simply the expression of certain laws ; a visible diagram of forces. We can, of course, if we like, inquire into the origin of the bubble and find causation there ; since the bubble is, let us say, the consequence of blowing air through a pipe against a flat film. But the shape of a bubble deformed by contact with a rigid surface is the resultant of the intrinsic laws of the system, on the one hand, and of the particular form of the surface, on the other ; it is thus a causal phenomenon with respect to the system. In Plateau's experiments, for example, the forms assumed by the soap-films were due to surface-tension *and* the particular shapes of the wire frames.

This distinction is particularly clear in the case of the soap-bubble, because this is an elastic system, and the deformations we have discussed occur only in the presence of some agent which interferes with the simple symmetry which the system has in more complete isolation. But in inelastic systems deformations are not only produced as the effects of contemporaneous interference, they also persist as consequences after the causal event.

If we turn now to living systems, we may consider that the embryo bird in its egg is the nearest approach to a perfectly isolated system we can find (no system in nature is, of course, completely isolated). Embryology can be considered from two points of view :

(a) Modern experimental embryology is chiefly concerned with the interaction of the parts of the embryo, and we are constantly reminded that it is a causal study. Some of its most interesting problems are causal problems ; for example, the problems of 'evocation' and 'competence,' that is, to what extent the actual moulding of a part (the 'formal cause'?) is a function of the cells taking part in the formation of the part, and to what extent it is a function of the inducing parts or substances. To take a familiar example, is it possible for some substance produced by the 'eye-cup' to induce a lens in any embryonal tissue, or must the cup itself be present ? Or must some particular tissue play the part of substrate ; and, if so, must we take this tissue at some particular time in its development ?¹ It is

¹ Cf. J. T. Needham, *Biochemistry and Morphogenesis*, Cambridge, 1942

THEORETICAL BIOLOGY

clear that these are causal questions, because they refer essentially to *interactions*.

(b) On the other hand, embryology can be considered as a nomological study, in so far as it is possible to treat the developing embryo as an isolated system having its own regularities ; and this clearly is justifiable, since the relations between the developing egg and its environment are of the most general kind. There are no correlates in the environment to the elaborate processes going on in the egg, and these processes, like the form of the soap-bubble floating in air, are explicable only in terms of the laws of the system itself. This statement has nothing to do with the question of whether the embryo is 'a mechanism' or 'a vital system.'

If the embryo, at least in the case of the embryo bird, can be studied as a virtually isolated system, with a minimum of reference to the environment, this is no longer the case for the adult organism as studied by the physiologist. But we notice that the physiologist tends to treat the organism as an elastic system having a 'normal' state of equilibrium to which it returns after every experimental (causal) intervention. This is not, of course, invariably so in physiological studies ; in experiments on growth, for example, the organism may suffer some permanent deformation. But on the whole, I think, it cannot be denied that the emphasis is on the reversible reactions of the system, which is hence treated as elastic.

When we turn to psychological studies, however, we are no longer able to exclude those phenomena in which the organism shows itself to be inelastic. The importance of this for our present purpose is that the psychologist is obliged to deal with causal phenomena in two ways, whereas the physiologist is only obliged to deal with them in one. The physiologist, certainly, cannot treat the system as the embryologist can, as having virtually no causal relations with its environment, but he can still consider the individual organism as a *type* : a healthy cat or dog is like an example of a machine turned out by a reliable firm, it has dependable and regular reactions to standard manipulations ; and although the method of investigation is largely causal, the final science of the organism which emerges is a nomology of the species. But, for the psychologist, not only is his method causal in the sense that he examines the organism by stimulating it, but the causal procedures do not leave the animal unaffected, the system itself is now changed as a consequence of its causal relations with its environment. Of course, the psychologist, if he is dealing with animals,

attempts to render his study nomological by giving a number of similar animals similar environments over long periods, but this is a precaution which the physiologist seldom needs to take. The difference may be brought out in this way: dogs commonly have the same reflexes, they do not commonly have the same memories.

Thus psychology is doubly a causal science, and except in unusually favourable circumstances, even an historical science. Of course, this is not the whole difference between psychology and physiology. I have made no mention of the fact that traces are retained by an animal otherwise than dents on putty or the evidence of wet and dry seasons in the rings of a tree. The mnemonic traces are not simply retained as they are received; they modify one another and become organised, and their organisation has its own nomology which seems to have no parallel, and consequently no explanation, in inorganic nature. For example, physical attractions and repulsions seem to be functions of space and time, whereas mnemonic traces seem to cluster together not only on spatial and temporal grounds, but on grounds of similarity.

It should be possible now to clarify the limitations of physiology as explanatory of psychology. It is easier to say that physiology cannot be reduced to chemistry than that psychology cannot be reduced to physiology, because one is clearer in one's mind as to what chemistry is than one is as to the meaning of the word 'physiology.' There is, of course, no reason why the physiologist should not use the training and testing techniques of the psychologist, and this has been done with great success. But this is not reducing psychology to physiology, it is simply combining the two disciplines. One thing can be said with certainty: in so far as physiology ignores the history, the 'biography' of the organism, it cannot fully explain the organism. There is also no doubt that physiology has in the past tended to abstract from this aspect of the organism, and rightly so.

How far physiology is likely to go in explaining the general laws of psychology it is very difficult to say, but we can at least say what would be necessary in order that explanation should be claimed. To explain physiologically the laws of memory, for example, it would be necessary not only to show how properties of the nervous system, identifiable by physiological techniques and not merely inferred from psychological, make the phenomena of memory *possible*; we must also show how the laws of these phenomena can be deduced from a purely physiological knowledge of the nervous system. If we could do this for all psychological laws, it would be possible to give a physiological

THEORETICAL BIOLOGY

explanation of the organism ; given certain environmental conditions not only at the time of the behaviour, but during the past life of the organism. But since it is manifest that organisms are not constrained by their structure to have the same experiences, the 'environmental co-ordinates' must always appear in the 'equation' ; and, if this is going beyond physiology, the explanation will not be wholly physiological.

5 Summary

I have attempted in this paper to consider some of the classical problems of explanation in biology in the light of a relative theory of causation. A cause is taken to be some factor to which we attribute the disturbance of some regularity. Living things are extremely complex systems in which the parts have causal relations *inter se* : the regular rhythm of the heart is disturbed by nervous stimulation ; the regularity of an undifferentiated field of competent ectoderm is disturbed by the presence of an optic vesicle near to some part of the field. But living things as relatively isolated systems can themselves be considered as complex regularities which are disturbed by environmental interference. Reaction with the environment can be considered as constituting a causal event only if there is some assignable disturbance of some regularity of the system : since the system is not totally isolated, it will have regular and constant relations with its environment which are necessary to its maintenance as a system ; thus the entry of oxygen into the system cannot as such be considered as a causal relation with the environment. On the other hand, episodic changes such as a temporary increase or diminution of oxygen in the immediate environment definitely disturb the rhythms of the organism and can thus be considered as causal. Among such episodic causal events are the reactions of environmental changes with the sense-organs.

Whatever can be deduced from the known properties of the system alone receives explanation, but not causal explanation.

Mnemic phenomena introduce causal events of peculiar interest. With respect to the organism they are causal, since there is nothing in the structure of the organism from which we can deduce the memories it acquires. Whereas in reflex activity the organism behaves as an elastic system, in memorising it behaves as an inelastic system. Mnemic traces, considered as imposed upon the structure common to a species, are consequences of causal events and therefore

J. S. WILKIE: THEORETICAL BIOLOGY

causal themselves with respect to the specific regularity of the species. But mnemonic traces do not remain indifferent to one another, as do the traces left in the wood of a tree by wet and dry seasons, they become organised *inter se*. Thus the behaviour of these traces has its own regularities which, since they are regularities, might be explained by a knowledge of the properties of the materials of the system (as cleavage patterns can be explained by molecular forces). This, however, appears improbable, since the physical language lacks the necessary concepts (constructs) under which the peculiar behaviour of mnemonic traces could be subsumed. Physical attractions all appear as functions of space and time, whereas mnemonic attractions appear as functions of space, time *and* similarity.

An attempt is made to justify teleological explanations, but it is not considered appropriate to treat these as causal explanations.

J. S. WILKIE

FUNDAMENTAL FEATURES OF CONTEMPORARY THEORY OF SCIENCE

I Introduction

IN a recent paper ¹ I have attempted to show that the classical conception of an apodeictic science, which was developed by Aristotle in his *Analytica posteriora*, no longer covers the present stage of scientific development. A number of divergent types of science have come into existence, most of which answer to some of Aristotle's postulates, but none of which satisfies all of them. Moreover, a number of disciplines to which Aristotle denied the name of science have come to obtain a strictly scientific character; this is especially the case with the great group of historical sciences which came of age during the nineteenth century.

Accordingly, in our times any attempt to construct an adequate theory of science has to face the danger of either indulging in irresponsible simplification or presenting a complete lack of coherence; this situation presents itself clearly to anyone who compares Mill's *Logic* (1843) with Wundt's *Logik* (1880-83) and who realises that since then no similar attempt at a comprehensive theory of science has been made. The strong preference for the monograph type of publication which is shown by serious specialists in this field is quite significant.

Therefore, if I am to explain some fundamental features of contemporary theory of science, I shall be forced to make a choice from this vast domain. I will concentrate upon the methodology of deductive science. My choice is based upon several considerations, four of which should be stated here: first, contemporary deductive sciences (logic and mathematics) come nearest to the type of science which Aristotle had in mind when he established his theory; secondly, these sciences derive by continuous development from sciences with which Aristotle was actually familiar; thirdly, the methodology of deductive science is at present in a more advanced stage of development than the methodology of other types of science; and fourthly, a methodology of deductive science affords an indispensable starting-point for the methodology of any other type of science.

¹ *The British Journal for the Philosophy of Science*, Vol. I, No. 1

As a matter of fact, the methodology of deductive science in itself constitutes already such a vast domain that here also I must be content to explain some of the more important topics.¹

2 *The Paradoxes of Logic and Mathematics*

It is well known that the development of modern logic and methodology has been strongly influenced by the discovery, since 1897, of a series of paradoxes which seemed to menace the foundations of the deductive sciences which, because of their apparently obvious character, had been taken till then to be entirely unshakable.

It will not be necessary here to set forth all these paradoxes. Again I will make a choice, and I will be content to state two paradoxes, which were known already in Antiquity, and to explain the methodological consequences to which their analysis has given rise. In my earlier paper I mentioned the fact that both of them were stated already by adherents of the School of Megara in their struggle against the doctrines defended by Plato and Aristotle.

I *The Liar Paradox*. 'If I say: "I am lying," do I lie, or do I speak the truth? You are lying! But if I lie, my statement was in agreement with the facts, so I spoke the truth. Then you were speaking the truth! But if I spoke the truth, my statement was contrary to the facts, so I was lying.' In this form, the paradox easily provokes too simplistic solutions and it does not offer an appropriate point of attack to a really thorough analysis. Therefore, we shall first give it a more suitable form.²

Let us suppose we are given a deductive theory T which has been formalised by means of the methods of symbolic logic; by this we mean, that the theory T is not formulated by means of ordinary language, but exclusively by means of formulae built up from arbitrarily chosen symbols in accordance with suitably chosen syntactical rules. We will suppose the theory T to contain the essentials of elementary arithmetic and we will assume that the usual symbolic

¹ For some comments on the methodology of physical science, I may refer to my paper 'Towards an up-to-date philosophy of the natural sciences' in *Methodos*, 1949, 1; for more details concerning the modern methodology of deductive science, to my *Fondements logiques des mathématiques*, Paris, 1950.

² For further details cf. K. Gödel, 'Ueber formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme', *Monatsh. Math. Phys.*, 1931, 38; A. Tarski, 'Der Wahrheitsbegriff in den formalisierten Sprachen', *Studia Philosophica*, 1933, 1; R. Carnap, *The Logical Syntax of Language*, London-New York, 1937.

CONTEMPORARY THEORY OF SCIENCE

apparatus A of elementary arithmetic forms part of the symbolic apparatus S by means of which the theory T is formalised. So the symbolic apparatus S is to include arithmetical formulae such as :

$$\begin{aligned}(x + y)^2 &= x^2 + y^2 \\ x(y + z) &= xy + xz \\ 3x + 5y &= 7\end{aligned}$$

etc.

It will, however, be clear that if our formalisation of the theory T is to be fully independent upon ordinary language, the symbolic apparatus S must include some elements which are not contained in the usual symbolic apparatus of elementary arithmetic. We must have at our disposal some symbols, for instance

$$\sim, \quad \vee, \quad \&, \quad \rightarrow, \quad \leftrightarrow,$$

which are to replace words and expressions such as 'not', 'or', 'and', 'if . . . , then . . .', and 'if . . . , then . . . , and conversely'; moreover, the symbolic apparatus S must contain symbols, for instance

$$(x), (y), (z), \dots, (Ex), (Ey), (Ez), \dots$$

which can replace the expressions 'for any x , we have . . .', 'for any y , we have . . .', 'for any z , we have . . .', . . . , 'there is an x such that . . .', 'there is a y such that . . .', 'there is a z such that . . .',

For instance, the arithmetical theorem : *the square of any odd number x is itself an odd number*, can be expressed as follows by means of the symbolic apparatus S :

$$(x)\{(Ey)[x = 2y + 1] \rightarrow (Ez)[x^2 = 2z + 1]\}.$$

By supposing the symbolic apparatus S to constitute an extension of the usual symbolic apparatus A of elementary arithmetic, we avoid the necessity of stating an elaborate system of syntactical rules for S. Incidentally it should be noticed that at present we are not interested in the problem of constructing an axiomatical basis for the derivation of the theorems of the theory T.

Moreover, it should be stressed that, besides the elements mentioned above, the symbolic apparatus S may contain additional elements $a, b, c, \dots, P, Q, R, \dots$, serving, for instance, to state geometrical theorems on physical laws, contained in the deductive theory T.

The formula

$$(Ey) [2x + y = 5] \leftrightarrow (x \leq 2)$$

will be said to contain the free variable x and the bound variable y ; the formula

$$x^2 + y^2 = (x - y)(x + y)$$

will be said to contain the free variables x and y , whereas the formula

$$(x)(\exists y)[xy \leq 3]$$

is said to contain the bound variables x and y ; finally, the formula

$$3 + 5 > 12$$

is said to contain neither free nor bound variables. A formula containing no free variables will be called a sentence ; any sentence will be either true or false ; among the formulae mentioned above we find a true sentence as well as a false one.

If n is a natural number, then by n^* we will denote the (indian) numeral by which n is usually denoted. It shall be a consequence of our assumptions that n^* will be contained in the symbolic apparatus S by means of which the theory T is formalised.

Now let $F(x, y, \dots, u)$ be a formula containing the free variables x, y, \dots, u , and no other ones ; then $F(x, y, \dots, u)$ will be called a valid formula, if and only if the sentence $(x)(y) \dots (u)F(x, y, \dots, u)$ is true. So

$$(x + y)^2 = 2xy + y^2$$

will be a valid formula, as

$$(x)(y)[(x + y)^2 = x^2 + 2xy + y^2]$$

is a true sentence.

Furthermore, let $F(x)$ be a formula containing the free variable x and no other ones. Then the natural number n will be said to fulfil the formula $F(x)$, if and only if the sentence $F(n^*)$ obtained by replacing the variable x everywhere by the numeral n^* is true. Obviously, $F(x)$ will be valid, if and only if it is fulfilled by any natural number n . For instance, the formula

$$(x + 2)^2 > 27$$

is fulfilled by the natural number 4, but it cannot be valid, as it is not fulfilled by the natural number 2.

Let α be any class of natural numbers and $F(x)$ be a formula containing the free variable x and no other ones ; then we can say that $F(x)$ constitutes a suitable definition for the class α , if and only if a natural number n is contained in the class α , if and only if n fulfils the formula $F(x)$. Conversely, the class α will be called definable within

CONTEMPORARY THEORY OF SCIENCE

the symbolic apparatus S , if and only if there is a formula $F(x)$ which constitutes a suitable definition for the class α .

The class α of all even numbers is obviously definable within the symbolic apparatus S , as the formula

$$(E\gamma)[x = 2\gamma]$$

constitutes a suitable definition for it. A formula can constitute a suitable definition for a given class, without being a more or less accurate translation of its current definition by means of ordinary language; for instance, the formula

$$(E\gamma)(Ez)[\gamma = (3x + 5)^2 \ \& \ \gamma = 2z + 1]$$

is easily seen to constitute a suitable—though somewhat eccentric—definition for the class of all even numbers.

Suppose the classes α and β to be definable within the symbolic apparatus S ; then their intersection—that is the class of all numbers contained both in the class α and the class β —will also be definable within the symbolic apparatus S ; indeed, suppose the formulae $F(x)$ and $G(x)$ to constitute suitable definitions for the classes α and β respectively; then the formula

$$F(x) \ \& \ G(x)$$

obviously constitutes a suitable definition for their intersection.

Now we introduce the method of the arithmetisation of meta-logic, which was first published by Gödel and independently discovered by Tarski. Let us draw up a list of all symbols, which are used in constructing the symbolic apparatus S ; this list may run as follows :
 $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -, =, <, >, \geq, \leq, (,), [,], \{, \}, \sim, \vee, \&, \rightarrow, \leftrightarrow, x, a, P, \gamma, b, Q, z, c, R, \dots$

We also draw up a list of prime numbers :

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, \dots$

Now suppose we are given a formula, for instance

$$2x < \gamma.$$

We consider the first symbol which it contains; this is the third in our list; we look for the third prime number and write it down :

5

Now we consider the second symbol contained in our formula ;

it is the twenty-ninth in our list ; we look for the twenty-ninth prime number and write down its square :

$$109^2$$

Then we consider the third symbol, which turns out to be the fourteenth on our list, so we write down the third power of the fourteenth prime :

$$43^3$$

As the fourth symbol in our formula is the thirty-second on our list, we write down :

$$131^4$$

Finally we take the product :

$$5 \cdot 109^2 \cdot 43^3 \cdot 131^4$$

which will be called the Gödel number of the formula under consideration. It will be clear that in this manner we can assign to any formula F contained in the symbolic apparatus S a univocally determined Gödel number $g(F)$. Of course, a given number n is not necessarily the Gödel number $g(F)$ of a formula contained in S ; however, if n is a Gödel number, then the corresponding formula F will be univocally determined by n .

Now we consider the class γ of all numbers n which satisfy the following conditions :

- (a) n is a Gödel number of a formula F contained in S ;
- (b) the formula F whose Gödel number is n contains the free variable x and no other ; we can therefore denote this formula by ' $F(x)$ ' instead of ' F ' ;
- (c) the sentence $F(n^*)$, obtained by replacing in the formula $F(x)$ the free variable x everywhere by the numeral n^* which denotes n , is false.

We will show that the class γ cannot be definable within the symbolic apparatus S . Indeed let us suppose the symbolic apparatus S to contain a formula $G(x)$ which constitutes a suitable definition for the class γ . Let g be the Gödel number of the formula $G(x)$. We ask whether g is contained in the class γ or not.

- (i) Suppose g to be contained in γ ; as γ is supposed to be defined by $G(x)$, the sentence $G(g^*)$ must be true. On the other hand as g is supposed to be contained in γ , it must satisfy the conditions (a)-(c) ; so on account of conditions (c) the sentence

CONTEMPORARY THEORY OF SCIENCE

$G(g^*)$ must be false; this supposition therefore leads to a contradiction.

- (ii) Suppose g not to be contained in γ ; as γ is supposed to be defined by $G(x)$, the sentence $G(g^*)$ must be false. On the other hand, as g is supposed not to be contained in γ , there must be one of the conditions (a)-(c) which is not satisfied by g ; as g clearly satisfies conditions (a) and (b), it cannot satisfy condition (c); so the sentence $G(g^*)$ must be true; this supposition therefore also leads to a contradiction.

It will be clear that there cannot be a formula $G(x)$ which constitutes a suitable definition for the class γ ; so the class γ cannot be definable within S .

It will be hardly necessary to dwell upon the analogy between the argument which shows the class γ not to be definable within S and the argument which gives rise to the liar paradox. Indeed, the sentence $G(g^*)$ can be taken to represent the statement which the liar is supposed to have made: the sentence $G(g^*)$ in a sense can be interpreted as stating its own falsehood. There remains, however, a slight but significant difference between the two cases. The liar's statement must be taken to exist; it is printed explicitly in this paper and the reader may read it aloud, if he should wish. The sentence $G(g^*)$ was merely supposed to exist; later it was proved that S cannot contain such a sentence. It follows that the structure of the symbolic apparatus S must in some respect be essentially different from the structure of ordinary language. On the other hand, the suppositions which have been stated concerning the symbolic apparatus S , namely

- (i) S includes the usual symbolic apparatus A of elementary arithmetic;
- (ii) S contains a certain number of symbols which can replace words and expressions such as 'not', 'or', etc.

do not seem to imply any essential difference between S and ordinary language; they rather seem to postulate a certain degree of similarity between S and ordinary language and indeed were meant to do so.

The essential supposition which implies the divergence of S with regard to ordinary language was introduced implicitly when we stated any sentence contained in S to be either true or false. The terms 'true' and 'false' with regard to sentences contained in S were used without any previous definition; but Tarski has shown that with regard to sentences contained in a symbolic apparatus S an

adequate definition of these terms can be given. Such a definition fills the gap we detected in our argument. It follows that with regard to sentences contained in ordinary language such a definition cannot be given; for otherwise we could apply our argument to ordinary language and show the liar's statement not to exist. Moreover, it follows that, if with regard to the sentences contained in a symbolic apparatus *S* an adequate definition of truth and falsehood can be given, then this definition cannot be reproduced within the symbolic apparatus *S* itself; for otherwise we would be able within the symbolic apparatus *S* to derive the liar paradox.

This last remark shows exactly what, from our present point of view, is wrong with ordinary language; for it will be clear that, if with regard to sentences contained in ordinary language an adequate definition could be given, then this definition could always be reproduced within ordinary language.¹ Notwithstanding Max Black's arguments to the contrary,² I am convinced that Tarski's analysis is of the greatest importance also from a philosophical point of view. Indeed, Tarski not only established a method which enables us to construct adequate definitions of truth and falsehood with regard to sentences contained in a symbolic apparatus such as *S*, but he also gave an accurate definition of the notion of an adequate definition of truth and falsehood in general, and he showed, on the basis of this definition, that an adequate definition of truth and falsehood with regard to sentences contained in ordinary language cannot be given.³

I will not dwell upon the results of Gödel (1931), A. Church (1936, 1937), and J. B. Rosser (1937), which are derived by means of arguments analogous to the one discussed above, nor can I enter upon the analysis of the paradoxes of Berry, Richard, and Zermelo-König, which give rise to other important insights of a methodological character.

II *The Larvatus Paradox*. 'Do you know your father? Yes! How is that possible? Suppose I show you a masked man and ask:

¹ For an elementary discussion of these matters, see A. Tarski, 'The semantic conception of truth', *Philosophy and Phenomenological Research*, 1944, 4, reprinted in: H. Feigl and W. Sellars, *Readings in Philosophical Analysis*, New York, 1949; an extensive discussion is also found in R. Carnap, *Meaning and Necessity*, Chicago, 1947.

² Max Black, 'The Semantic Definition of Truth', *Analysis*, 1948, 8.

³ Therefore, Black's argument to this effect (*loc. cit.* p. 56) seems to be superfluous and irrelevant; Black seems to overlook that Tarski's result also impedes the realisation of his desideratum of a general criterion of truth (*loc. cit.* p. 60).

CONTEMPORARY THEORY OF SCIENCE

“Do you know this man?”—what will be your answer? That I don’t know him, of course! But now that man happens to be your father! So, if you don’t know that man, you don’t know your own father.’

This paradox may be considered as the prototype of the paradoxes of denotation and of analysis which, though certainly less profound than the liar paradox, have recently been discussed in much detail.¹

Already in our discussion of the liar paradox we incidentally had to take into account the fundamental distinction between use and mention of a symbol. Let us consider the two sentences:

Charlemagne was a King of the Franks.

Charlemagne contains eleven letters.

The first sentence refers to a certain historical personality whereas the second one refers to a certain word; this difference, however, is not explicitly expressed. In ordinary discourse, this lack of accuracy is no obstacle to mutual understanding; in logical analysis, it turns out to be a real hindrance.

In our discussion of the liar paradox, for instance, it was necessary explicitly to state that any natural number n should have a notation n^* in S ; and this is, of course, only possible if, in referring to a natural number n and to the corresponding notation n^* , we call them by different names. But in the two sentences considered as an example, one and the same word was first used in order to refer to a certain historical personality, and then used in order to state one of its own properties; in the second case, instead of using the word itself, we should have referred to it by using a name for it. In general, if we want to state properties of words or symbols, then we should refer to these words or symbols by using names for them.

A well-known and practical device for satisfying this requirement of logical accuracy consists in referring to a word or symbol by using as a name for it the same word or symbol in quotation-marks. If we adopt this convention, we should write

Charlemagne was a King of the Franks.

‘Charlemagne’ contains eleven letters.

¹ A very able and thorough study has been devoted to these paradoxes by R. Carnap, *Meaning and Necessity*, Chicago, 1947, where also references to earlier publications will be found. In our times, studies on this topic were initiated by Frege (1892) and Russell (1905). Cf. also Tarski’s paper which was mentioned on p. 292, n. 2.

As an obvious consequence, we also should write

‘Charlemagne’ denotes a King of the Franks
 By ‘Charlemagne’ we refer to a King of the Franks
 ‘Charlemagne’ derives from ‘Carolus Magnus’

and even

By “‘Charlemagne’” we refer to the name of one of the
 Kings of the Franks.

Other examples are :

‘Prince of Wales’ is the traditional title of the eldest son of the
 King of England.

The introduction of indian numerals ‘1’, ‘2’, ‘3’, . . .
 instead of roman numerals ‘I’, ‘II’, ‘III’, . . . as a notation
 for natural numbers 1, 2, 3, . . . constitutes an improvement
 from a theoretical as well as a practical point of view.

In algebra it is usual in order to simplify notations, to write
 ‘ $abc + efg$ ’ instead of ‘ $(a \times b \times c) + (e \times f \times g)$ ’.

Now this convention in its turn gives rise to the so-called paradox
 of denotation (or antinomy of the name-relation). Let us denote by
 $s[u]$ the number of elementary symbols contained in a symbol ‘ u ’.
 Then, for instance, we have

$$s[23] = 2, \quad s[\log 5] = 4, \quad s[(x + y)^2] = 6, \text{ etc.}$$

Now compare the two inferences :

$\log 10 = 1$	$s[10] = 2$
$1 + 2 + 3 + 4 = 10$	$1 + 2 + 3 + 4 = 10$
$\log (1 + 2 + 3 + 4) = 1$	$s[1 + 2 + 3 + 4] = 2$

The first inference belongs to a familiar type, which is frequently
 applied in mathematics. The second inference also appears to belong
 to this type, but the conclusion to which it leads is manifestly false, as
 on account of our definition we find $s[1 + 2 + 3 + 4] = 7$.

In order to deal with this paradox, we observe that the expression
 ‘ $\log 10$ ’ originates from the expression ‘ $\log x$ ’ by substituting
 ‘10’ for ‘ x ’, whereas the expression ‘ $\log x$ ’ serves to characterise a
 certain function of the variable x ; let us use the name ‘matrix’ for
 any expression which in a similar manner characterises a function of one
 or more variables. Then it will be clear that we always will obtain a
 stringent inference, if we replace the matrix ‘ $\log x$ ’ by any other

CONTEMPORARY THEORY OF SCIENCE

matrix. So the expression ' $s[x]$ ' cannot be a matrix ; this is no wonder as the expression " x " itself cannot be taken to be a matrix, that is, to characterise a function of the variable u ; indeed, from $u = v$, it does not follow that ' u ' = ' v '.

So the only conclusion from our analysis of the paradox of denotation seems to be that an expression which contains a letter ' x ' or ' u ' need not always be a matrix. This elementary insight enables us at once to disentangle the paradox of analysis too. Let us consider the geometrical theorem :

The class of equiangular triangles is identical with the class of equilateral triangles.

On account of this theorem, if in a sentence referring to the class of equiangular triangles we replace the words 'equiangular triangles' by 'equilateral triangles', we will get a sentence which is logically equivalent to the original sentence. Now if we apply this operation to the theorem under consideration, we obtain the theorem :

The class of equilateral triangles is identical with the class of equilateral triangles ;

which consequently is logically equivalent to the original one. On the other hand, the first theorem being non-trivial, the second theorem which is logically equivalent to it, should be expected to be non-trivial too, whereas it is notoriously trivial. The fallacy underlying this paradox of analysis stands out immediately if we give it the form of an inference ; for the sake of brevity, let α be the class of equiangular triangles, and β the class of equilateral triangles ; then we obtain the inference :

the theorem ' $\alpha = \beta$ ' is non-trivial

$$\alpha = \beta$$

the theorem ' $\beta = \beta$ ' is non-trivial.

Again the mistake has been made of considering the expression " $\alpha = \beta$ " as a matrix, characterising a certain function of α and β .

It is interesting to observe that a similar fallacy underlies the opinion according to which the theorems of pure mathematics, being logically equivalent to ' $I = I$ ', are mere tautologies.

Let us now, for a moment, return to the larvatus paradox ; we can give this paradox the form of the following inference :

EVERT W. BETH: THEORY OF SCIENCE

Aristotle knows the name of his father

The masked man is identical with his father

Aristotle knows the name of the masked man.

Again the fallacy has been committed of considering the expression 'the name of x ' as a matrix.

Conclusion

This summary of some recent developments in the methodology of deductive science may suffice to show that Aristotle's conception of pure science as being the fruit of a purely contemplative attitude of the mind, which should enable us to grasp those obvious truths to which an active way of living might make us blind, fails already with regard to deductive science. Even in the domain of pure logic, which seems so remote from practical life, those obvious insights which offer themselves so naturally to the relaxed intellect turn out to be in no way reliable; conceptions which can sustain a thorough critical examination can be conquered only at the expense of strenuous work. This holds true even for the most fundamental and seemingly the most elementary methodological principles of the deductive sciences.

This criticism of Aristotle's conceptions should, however, not induce us to overlook his exceptional merits in the developing of deductive science and thereby of science in general. It is incontestable that his errors have seriously hampered the expansion of science, but on the other hand it is hardly probable that without his contributions, which I do not hesitate to qualify as worthy of a genius, the weak plant of early Greek logic and mathematics ever should have survived.

EVERT W. BETH

ERRATA

Professor P. Bernays has kindly drawn my attention to the fact that the assignment of Gödel numbers to formulas described on page 295 of volume I of this *Journal*, is not correct. Under this assignment, different formulas may obtain the same Gödel number and therefore the argument given on page 296 under (i) is not cogent. The correct assignment would give, for the formula considered as an example, the Gödel number

$$2^3 \cdot 3^{29} \cdot 5^{14} \cdot 7^{32}$$

The correction to be made on page 294, l. 20 is obvious.

E. W. BETH

FUNDAMENTAL PHYSICAL THEORY

AN INTERPRETATION OF THE PRESENT POSITION OF THE THEORY OF PARTICLES

1 *Introduction*

FUNDAMENTAL theoretical physics is now in an exceptionally interesting situation :

(a) Since the development of quantum mechanics in 1925-1927 there has been no major advance in basic theory. It seems that a fresh start may be necessary involving radical changes in present physical ideas.

(b) During this period our knowledge of nuclear properties and of cosmic rays has been greatly extended and the phenomenon of particle transformations has been discovered. These properties appear to lie beyond the scope of established methods, the provisional theories now in use being unsatisfactory.

(c) Fundamental theory has become very complex, and if simplicity is to be restored entirely new methods are probably required.

(d) The theory of elementary particles holds a key position in fundamental physics, and it is widely felt that the next major step must be the establishment of a comprehensive theory of the interactions and transformations of simple and compound particles, probably based not on dynamical concepts and analogies, but on some novel physical principle.

A fundamental advance in our understanding of particles can come about as the result of further *experimental* work (e.g. on nuclei, mesons, and cosmic rays), of improvements in existing *mathematical* methods (e.g. use of non-linear differential equations), of fresh *theoretical* ideas (e.g. a theory of measurement taking into account all actual restrictions), or of combinations of these. Experiment and mathematics are always necessary, but there are times when theoretical enquiries, such as the examination of fundamental assumptions, may help to guide new mathematical formulations, as was shown by Einstein's approach to relativity, and Heisenberg's to quantum theory. Thus timely studies in the philosophy of a science, if directed towards outstanding difficulties, can further its advance.

This paper offers an interpretation of the present position of the theory of the elementary particles in the light of a theoretical research programme¹ drawn up before these new phenomena were explored. The aim was 'to discover a more general limitation of the classical frame which reduces in different special cases to the relativistic and quantum limitations associated with c and h respectively and at the same time accounts for the appearance of the constant e . This new method is to be reached by a final elimination from fundamental theory of the assumption of a four-dimensional world describable in terms of four metrical coordinates.' This programme is still unachieved; indeed none of the unsolved problems listed for attack have yet been satisfactorily treated. Nevertheless the methods then proposed lead to a constructive interpretation of the present position. This interpretation can be developed without detailed reference to the mathematics of atomic or nuclear theories because it is based on the assumption that at the present time we must pay special attention to the *most directly observed quantities*. What we regard as 'directly observable' at any time depends on the theories in use, and today there is no satisfactory fundamental theory. We are, therefore, compelled to build, as far as we can, by using only the *most direct* measurements.

The validity of quantum theory over wide fields is well established. But its standard methods are probably inadequate to deal with certain problems which fundamental theory now faces, and the first step towards an advance may be to overcome the mental climate of 1926 and to open up new vistas. This paper therefore concentrates on some of the most directly observable facts which appear now to require closer attention than they were given in the general theory of quantum mechanics.

The following notation will be used for the five primary dimensional constants of atomic theory :

- c the velocity of light in vacuo
- e the electronic charge
- h Planck's constant of action
- m the rest-mass of the electron
- M the rest-mass of the proton

and for the two pure number constants derivable from them :

$\alpha = 2\pi e^2/hc = 1/137.0(3)...$ the fine-structure constant

$\beta = M/m = 183(6)...$ the ratio of proton to electron mass.

¹ L. L. Whyte, *Critique of Physics*, London, 1931, p. 142

FUNDAMENTAL PHYSICAL THEORY

Possible discrepancies or secular changes in any of these quantities and their connections with gravitation and cosmology will be neglected. Other masses (such as those of unstable particles) and other pure number constants (such as the ratios of meson to electron masses, and the expression $2\pi g^2/hc$ in Yukawa's theory of short range forces) will be treated as of secondary importance.

2 The Changing Concept of the 'Particle'

During the early years of this century physical theory made use of only one kind of elementary particle, the *electron*, or small negative charge exerting a central force, whose position in space could be specified by a mathematical point, associated with an inertial mass, and possessing individuality. The only characteristic parameters associated with the electron were its charge e , and mass m .

Today the position is very different :

(a) Some ten or more different types of particle (in a generalised sense) are now in use : the *photon* ; *electron* and *positron* (two states of one light particle) ; *neutrino* ; several *mesons* ; and the *proton* and *neutron* (states of one massive particle, the *nucleon*).

(b) These particles are not permanent, being created and annihilated, or transformed into each other, under suitable conditions. The property of individuality has gone, the three-dimensional path and the time coordinate of 'an individual particle' not being traceable through a complex scattering process. Nucleons alone may constitute an exception and enjoy permanence, though it has been suggested that nucleonic matter may be created during the history of the universe.

(c) Under certain experimental conditions the property of location at a series of successive points in space disappears, and the phenomenon displays the extended spatial pattern of a stationary wave-field. The time coordinate which plays an essential role in the kinematics of particles has largely faded out in the wave representation of elementary processes in small regions. The 'particles' and 'waves' appear to be to some extent products of circumstance, i.e. the consequences of particular experimental procedures, interpreted in terms of current theories.

(d) The parameters associated with these particles are : $\pm e$; m , M , and the series of meson rest-masses ; c ; h ; spin values and magnetic moments ; and life-periods. The masses vary with velocity and are not always additive, fractional differences being convertible into energy.

(e) The nuclear particles obey obscure short-range laws, showing effects equivalent to 'non-central forces,' though it is possible that intra-nuclear processes may involve *n-term* relations not reducible to forces between *pairs* of point-localised entities.

(f) Other less definite changes in the conception of particles result from the transformation in the general character of physical theory during this century from exact causal to statistical description, and from the assumption of the existence of physical entities possessing measurable properties to reliance on the procedures by which physical numbers are obtained and the conditions under which this is possible.

(g) In much of the most recent work, for example in the relativistic theory of the finite electron, there is evident a need to overcome the dualism of material particle and force-field, and to deal directly with the measurable parameters involved in the interactions of particles.

Yet in spite of this dissipation of the classical particle, the discrete, localised aspect of physical phenomena is as evident as ever it was ; indeed more so. The crux of current fundamental physics is displayed in the fact that during the very period when fundamental theory seems to have discarded the last vestiges of the classical mechanical particle, experiment has made directly visible the tracks of single ultimate particles. The general particle theory of the future, or some theory equivalent to a theory of particles, has to resolve this paradox.

One direction of advance towards this objective is already clear. Even in classical electron theory, which put 'the radius of the electron' at about 10^{-13} cm, it was evident that the concept of the point-electron would prove inadequate at very small distances, though it could provide a good approximation towards atomic structure.¹ More recently Bohr,² Born³ and Heisenberg⁴ have stressed the fact that the use of the idea of point charges in the theory of atomic or molecular structure is justified only by the large size of atoms (of the order of $h^2/4\pi^2me^2$) compared with e^2/mc^2 , i.e. by the small value of α compared with unity. Thus the utility of the point-particle concept of the electron is bound up with the smallness of α . The point-electron representation corresponds to the first term in a series providing successive approximations to the actual phenomenon, and if the distances involved are

¹ E.g. see O. W. Richardson, *Electron Theory of Matter*, Cambridge, 1916, p. 555

² N. Bohr, *J. Chem. Soc.*, 1932, p. 349

³ M. Born, *Proc. Ind. Acad. Sci.*, II, 1935, 6, A, 533

⁴ W. Heisenberg, *Two Lectures*, Cambridge, 1949

FUNDAMENTAL PHYSICAL THEORY

very small (say of the order e^2/mc^2) this term no longer even provides a useful first approximation. One difficulty is that dynamical analogies which are known to break down in a certain region cannot even be used to provide arguments indicating exactly where that region begins. This is the reason for some ambiguity in the literature regarding the exact stage at which the point-electron breaks down, i.e. fails to provide a given accuracy of representation.

In this situation reliance has to be placed not on the particles, but on the measurable parameters associated with elementary processes. These parameters are no longer to be regarded as representing the magnitude of some physical property of an isolated entity—for we can never know anything about isolated entities—but as representing components of the processes which still have to be described as the *interactions* of simple or compound particles.

Eddington¹ defined a particle as a *conceptual carrier of a set of variates* (i.e. the occupant of a state defined by a set of variates). This still puts the emphasis on 'the single particle' rather than on the interactions and transformations of particles. But no theory based on the concept of particles can provide a satisfactory theory of their transformations, since that requires the representation of the different particles as special cases of some more general principle. If we discard the term 'elementary particle' as inadequate for fundamental theory, and substitute the term 'elementary observable parameter,' then we can say that the aim of 'particle theory' is a *theory of the transformations of elementary observable parameters*. As Bhabha has said, the elementary particles are but the transitory embodiments of the metamorphoses of the invariants (such as mass-energy, momentum, and charge) which are the permanent things in nature. It seems that, at least in considering certain fundamental problems, we have to learn to forget the particles, and to concentrate on the directly observable parameters and their transformations, until some new physical principle is discovered.

This is the purport of Heisenberg's² scattering matrix, S. Heisenberg sought to overcome the divergence difficulties arising from the presence of a universal length in the Hamiltonians of current field theory by retaining only those quantities which are 'directly observable' in a generalised scattering process (these are assumed to be

¹ A. S. Eddington, *Fundamental Theory*, Cambridge, 1946

² W. Heisenberg, *Zeits. f. Phys.*, 1943, **120**, 513, 673

energies of stationary states, collision cross-sections, and parameters of free particles). The use of dynamical representations (Hamiltonians) based on classical analogies may be only asymptotically valid at large distances from the scatterer (i.e. as $r \rightarrow \infty$, or $r \gg e^2/mc^2$), since exact measurement of distances and localisation of field quantities is probably impossible within regions comparable to the size of the nucleus. It is therefore necessary to construct the scattering matrix, or transformation operator S (already used by Wheeler), which converts the incoming set of Schrödinger, Maxwell, and/or other fields into the outgoing set of fields, both being treated together as one 'stationary state.'

The matrix, S , is at this stage only an abstract mathematical idea, since where $r \sim e^2/mc^2$ S is not associated with localised fields and therefore cannot derive physical significance from classical dynamical analogies. But it can be regarded as a special form of the relativistic density matrix, ρ , of Schrödinger's theory. S links the incoming and outgoing probabilities, and determines with what outgoing wave a given incoming wave must be supplemented in order to satisfy the scattering laws. Though approximations to S covering simple scattering processes can be obtained from known principles, Heisenberg considers that the general form of S determining all interactions and transformations (including high energy processes involving multiple particle transformations) must ultimately be derived from some new physical principle not using localised field functions. The difficulties of current theory can probably be overcome only by treating all elementary particles and all their potential transformations together. If so, what is required is a general theory of the transformations of elementary observable parameters, in which all changes of state, transitions, and transformations, whether of simple or compound 'particles,' are treated simultaneously. In such a theory the present sharp distinction between intra-nuclear and extra-nuclear processes, e.g. between radio-activity and radiative transitions, would appear much less fundamental. It is unlikely that this far-going result can be achieved by a theory based on classical analogies.

Heisenberg's proposal carries physical theory potentially beyond dynamics and dynamical analogies into a new and relatively unexplored region, in which the basic laws may not involve x , y , z , and t , at least as now used. The new physical principle which is needed to give form and content to S and to provide a comprehensive theory of the nucleus

FUNDAMENTAL PHYSICAL THEORY

is unknown. So the suggestion is still an empty frame, and even the meaning of 'observable quantities' is itself uncertain when so radical a theoretical transformation is in view.

Nevertheless one feature of the proposed method is already becoming clear. Dyson¹ has thrown light on the relation of the old quantised wave-fields with localisable interactions based on the use of Hamiltonian dynamical analogies to the new scheme of asymptotic directly observable quantities subject to the transformation matrix S . The former represents the picture established by an *ideal* observer whose possible measurements are imagined to be restricted only by the finitude of c and of h , and the latter the picture obtained by a *real* observer, whose measurements are subject to the additional restrictions set by the fact that his apparatus is composed of ultimate particles, i.e. by the finitude of e , m , and of α . The ideal observer is a physicist trained in 1927, the real observer one schooled in all the actual limiting conditions of measurement as known in 1949 (including the 'particle structure' of his measuring instruments), which for example prevent the determination of the exact strength and exact coordinates of a single field vector, undisturbed by its relations with other fields. Thus the real observer is compelled to recognise that the localised wave-fields of contemporary physics represent only *asymptotic components* (as $r \rightarrow \infty$), or first approximations, of *something else*. The old conception of the field is no longer satisfactory, because the determination of the field strength at a point depends on the character of the test-object used, i.e. on its charge, mass, etc.² and a new complementarity appears between measurements of the field magnitudes and of the x , y , z , t , of the test-object.

Heisenberg treats S as an abstract mathematical operator which links actual physical entities, such as electrons, photons, mesons, etc. But this may be misleading. For if his trend of thought is right, it is the 'electrons,' 'photons,' etc., which are special limiting cases or asymptotic mathematical components of something else, and that something else is the basic physical phenomenon. Thus electrons and photons should be linked, not merely by an abstract mathematical transformation, but by interpreting them as special components of a single underlying physical phenomenon, or *primary field*, whose mathematico-physical character has still to be discovered.

¹ F. J. Dyson, *Phys. Rev.*, 1949, **75**, 486, 1736

² M. Markow, *J. Phys. U.S.S.R.*, 1940, **2**, 453

On this view the various particles are not to be regarded as 'independent' entities capable of 'interacting', but as partially separable components of a primary field, where the term 'field' does not imply a force-field, like those of classical theory, but is in some sense a direct representation of the total phenomenon. The laws of the primary field will not be expressed in terms of particles, and will use only real observables, taking into account all the restrictions set by c , h , e , m , M , etc., including α and β . These primary laws must determine all the phenomena of particle physics, the different elementary particles being represented by the sets of parameters characterising particular, partially separable, asymptotically localisable ($r \gg e^2/mc^2$) components of the primary field.

The fact that these parameters characterise components which are only asymptotically separable (as $r \rightarrow \infty$) and are not exact quantities of the primary field, is evidenced empirically in the phenomena described as the 'interactions of the particles.' Thus the emphasis has to be shifted from the 'individual particles' to the 'sets of asymptotically separable parameters.' We must no longer speak of the *particle structure* of a complex system, but concern ourselves with the *parametric structure* of the changing phenomenon, of which the stationary or stable system represents only one aspect. Though the mathematico-physical character of the primary field is unknown, the conception of the primary field suggests that theoretical enquiry must be pursued in a direction which can already be clearly defined, namely, from entities localised in (relativistic) space-time displaying 2-term interactions (forces), towards a representation, not using four coordinates, of the changing parametric structure of systems involving n -term relations, where n is variable (because particles appear and disappear) and is in general > 2 .

The search for the laws of the primary field must be guided by closer attention to those measurements of elementary processes which are *most direct*, i.e. involve least explicit or implicit assumptions and the minimum of theory to justify or explain them. Now if we forget for a moment the sophistications of current physical theory, and use only stubborn common sense, we are left with little doubt that the most direct class of measurements (which is also far the most extensive) in atomic physics are *determinations of lengths or distances*. These may be measurements of wave-lengths, curvatures of orbits, atomic radii, lattice-constants, etc., and ranges. This class not only covers a very

FUNDAMENTAL PHYSICAL THEORY

high proportion of all the measurements on which electron theory, atom theory, and quantum theory are based, but includes nearly all measurements in atomic physics other than those interpreted in classical theory as statistical results of large numbers of elementary processes (e.g. life-periods of atoms, and energy density in black-body radiation, either total or at λ_m ; see section 3). Determinations of lengths are thus both the most extensive and the most direct class of measurements in particle physics.

The only other types of measurement which might conceivably compete are those of mass, and of time. But particle masses are inferences from actual observations of lengths, and determinations of time periods are relatively few, are mostly statistical in character, and have played a minor role in the development of atomic and quantum theory. (An electron physicist put it thus: 'We don't use a stop-watch in our lab.') For example, the most direct determination of electron velocity depends on balancing the transit time over a finite path against the phase-change of an oscillatory electromagnetic field of known frequency. In all other methods a spatial relation is measured, usually the curvature of a line in three-dimensional space.

Moreover it is the presence of universal lengths directly related to observable lengths (not times) which is the source of the recent difficulties in field theory. We have seen that theoretical enquiry should as far as possible concentrate on current difficulties. We propose therefore to examine afresh the *directly observable lengths* which appear in quantum physics.

3 *The Fundamental Lengths of Atomic Physics*

A basic difference between classical and quantum physics is that the former deals mainly with *scale-free* and the latter with *scale-fixed* phenomena^{1, 2} (the terms come from Eddington³). Classical physics is concerned with quantities whose laws do not involve a natural unit of linear scale, while quantum physics using h deals with quantities which are directly related to fundamental structures of definite spatial magnitude.

In Newtonian gravitational dynamics and Coulomb electron

¹ Whyte, *Critique of Physics*, p. 12

² Whyte, *Zeits. f. Phys.*, 1929, **56**, 809; 1930, **61**, 274

³ Eddington, *Fundamental Theory*, p. 16

dynamics under the inverse square law any possible (closed) orbit can be transformed into an infinite number of geometrically similar orbits of different sizes, in which the corresponding times are proportional to the $3/2$ th power of the linear scale of the orbit (Kepler's Third Law). Without this property the use of a single constant for gravitational orbits, and of e/m for all electronic orbits, would not be possible. Similarly in Maxwellian field theory any solution of the field equations remains a solution if the linear scale and the time periods are both multiplied by the same arbitrary factor, and this property alone permits the use of a single constant c . This 'principle of dynamical similarity'¹ holds in each of the separate realms of classical theory, each of which is characterised by a dimensional constant involving both L and T , and none of which contains in its general laws a constant with the dimensions of a length. In this sense classical physics was scale-free, and classical theories were therefore necessarily incapable of providing a theory of the structure of matter and radiation, these being scale-fixed phenomena.

As far back as 1763 it was evident to Boscovitch that any theory which assumes that solid bodies are constituted of a finite number of atoms must include a unique length, and he therefore used a law of force changing from repulsion ($r < r_0$) to attraction ($r > r_0$) at a distance, r_0 , which determines the size of all bodies composed of a definite number of atoms. It is impossible for any theory to represent the definite sizes of nuclei, atoms, molecules, crystal lattices, characteristic wave-lengths, etc., relative to the selected unit of length, without using at least one fundamental constant of the dimensions of a length (or its equivalent). That is a theoretical necessity, and on the empirical side by 1805 Young had already made a reasonable estimate of the size of molecules. So it was open to anyone to realise, even prior to the experimental discovery of e and h , that Newtonian dynamics and Maxwellian field theory could not possibly be adequate for a comprehensive theory of phenomena already known.

We are wise after the event; it seems that only Larmor² was wise before it. He published in 1900 a demonstration (written in 1898) that 'the definiteness of scale of the molecules of material systems thus precludes the possibility of their being constituted of singularities of a uniform continuum (whose laws obey dynamical

¹ For references see Whyte, *Critique of Physics*, p. 148

² J. Larmor, *Aether and Matter*, Cambridge, 1900, pp. 189-192, 233

FUNDAMENTAL PHYSICAL THEORY

similarity) with nuclei undistinguishable from mathematical points.' Larmor saw the far-reaching implications of 'this deficiency of definite scale' in the physical theories of the time, and thus came near to predicting the discovery of a new constant.

But when Planck introduced the additional constant which (with e and m) was indispensable to any theory of atomic structure, he was led by the experimental context of his discovery and the structure of physical theory at the time to give it the dimensions, not of a length, but of action. Thirty years were to pass before it was suggested^{1, 2} on general grounds that *the primary significance of h is to provide (with the other constants) the standards of linear scale appropriate to the description of the structure of matter and of radiation.*

The following table shows how the combinations $h^3c/4\pi^2me^4$, $h^2/4\pi^2me^2$, and h/mc are used in the theoretical representation of the structure of matter and radiation, and includes also e^2/mc^2 , the status of which in physical theory is less clear.

Four Fundamental Lengths of Atomic Physics

(Involving c , e , m , and h only ; ratios involving powers of α)

$h^3c/4\pi^2me^4$ $\cdot 914 \cdot 10^{-5}$ cm. (= $1/R$). May be called the *Rydberg wave-length*. Contains all four constants. Determines the scale of all wave-lengths in characteristic line spectra. May be regarded as directly measured.

$h^2/4\pi^2me^2$ $\cdot 529 \cdot 10^{-8}$ cm. *Radius of 1st circular Bohr orbit of H with fixed nucleus*. Does not involve c . Determines all stable atomic and molecular sizes, crystal lattices, etc. This may be regarded as the most directly measured fundamental length, since it determines the structure of all rigid bodies and measuring rods. The point-electron provides a good first approximation in regions of this order.

h/mc $2 \cdot 43 \cdot 10^{-10}$ cm. *Compton wave-length for electron*. Does not involve e . Determines wave-lengths (or their differences) in the fine-structure of line spectra, in photon-electron energy interchanges, and in Compton shifts. This length is characteristic of quantum mechanics, and Bohr's Correspondence Principle is normally reliable down to regions of this order. Lengths of this order

¹ Whyte, *Critique of Physics*, p. 21

² Whyte, *Zeits. f. Phys.*, 1929, **56**, 809 ; 1930, **61**, 274

and upwards may be regarded as directly measurable, since based on reliable analogies.

e^2/mc^2 $2.82 \cdot 10^{-13}$ cm. *Classical electron radius* (approx.). Does not contain h , and is not directly involved in quantum theory or electrodynamics. The inferred nuclear radii are of this order, this alone justifying the sharp separation of intra- and extra-nuclear properties. The point-electron and all dynamical analogies break down in regions of this order, radiative reactions and ponderomotive forces being inseparable here. Since all existing physical arguments rest on classical analogies which are invalid here, even the consistency of the indirect determinations of nuclear radii does not prove them valid. All phenomena apparently of this scale may ultimately require reinterpretation in a new theory.

The importance of the need for definiteness of scale in a structural theory, as intuitively evident to Boscovitch, explicitly suggested by Larmor, and now proposed as the basis of a future re-interpretation of h , lies in its generality and direct relation to observation. It follows from Earnshaw's Theorem ('a charged body placed in an electric field of force cannot rest in stable equilibrium under the influences of the electric forces alone') that no stable configurations of classical electrons are possible in which the electrons are either at rest or oscillating about positions of equilibrium, so that they must describe orbits and therefore radiate. Bohr¹ was consequently led to emphasise the role of h in determining the stability of particular electronic orbits and providing an appropriate fundamental length. But these arguments depend on special electrical and dynamical principles, whereas it can be seen at once that *no* physical theory of any kind explicitly using measurements of lengths can represent the definiteness of scale of fundamental structures and characteristic wave lengths, unless it contains an appropriate constant with the dimensions of a length, or its equivalent. The generality of this fact, and its direct foundation in spatial measurements, suggests that since h does in fact supply, with e , m , and c , appropriate constants, we have here a clue to the most general significance of h and to its role in a future theory.

After the discovery of the electron the combination e^2/mc^2 was

¹ N. Bohr, *Phil. Mag.*, 1913, 26, 1

FUNDAMENTAL PHYSICAL THEORY

potentially available to give definiteness of scale to fundamental theory, so that h might not have been necessary in addition. But e^2/mc^2 is of the order of 10^{-13} cm, whereas atomic radii are of the order $\cdot 5 \cdot 10^{-8}$ cm. So electrons could be treated as point charges in the theory of atomic structure, and e^2/mc^2 did not provide an appropriate length for atomic theory.¹ Thus Planck's constant was necessary.

Moreover an examination of the empirical data from which h can be calculated shows that the role of h in relation to fundamental phenomena (not involving temperature effects) is to represent *observable lengths* :

(a) *Fixed characteristic lengths* (wave-lengths of line spectra including fine- and hyper-fine structure, the wave-length determining Compton shifts, atomic and molecular sizes, crystal lattices).

(b) *Correlation of variable wave-lengths* with curvatures of electron paths or electron energies (electron energy = hc/λ , photon ; electron $p = h/\lambda$, electron).

(c) h also fixes the scale of the *quantum indeterminacy*, due to interactions, in the process of the measurement of the above wave-lengths.

This analysis shows that the role of h in fundamental, non-thermal, theory is to permit the representation of directly observable lengths connected with the structure of matter and radiation. But this fact has been concealed because Planck introduced h in connection with a statistical theory of temperature radiation, and because the dynamical concept of *energy* rather than the structural concept of *length* has dominated physical theory throughout this century.

The proposed interpretation of h is not arbitrary, because it is based on directly observable facts, nor is it trivial, because, as we shall see, it opens up new realms of theoretical enquiry. All other interpretations of h depend on the structure, not necessarily of basic physical phenomena, but of particular physical theories, and are therefore more liable to be upset in the future. Moreover the interpretation given to h by Planck and Einstein during the period 1900-1905, as essentially a quantum of dynamical *action* leading to quanta of radiant *energy*, cannot be left unmodified if physical theory is ultimately to interpret e and h as dual expressions of some underlying common principle (see section 4). There is therefore good ground for exploring the possibility that the deeper, i.e. more general and more permanent, interpretation of h is that anticipated by Larmor before h was discovered

¹ Cf. J. J. Thomson, *Electron in Chemistry*, Philadelphia, 1923

experimentally in a special field and given a particular, unduly restricted interpretation. On this view what the quantum physicists have discovered is not correctly described as the existence of *quanta* of action or energy, for that inference depends on physical concepts now proving inadequate, but is *the definiteness of linear scale inherent in all elementary physical phenomena*.

This interpretation of h is not generally accepted, and cannot be regarded as fully established until it has been made the basis of a successful theory. Eddington noted that all quantum phenomena involving h are scale-fixed, and the importance of fundamental lengths in atomic theory has been widely recognised since 1929. For example the role either of e^2/mc^2 or of fundamental lengths involving h has been stressed by Fürth,¹ Bohr,² Born,^{3, 4} Podolsky,⁵ Heisenberg,^{6, 8} Flint,⁹ March,¹⁰ and others. From 1930 onward many workers have considered theories of 'discrete space-time', or of a 'minimum observable length and proper-time,' but these appear to have led to no constructive results. In 1936 Heisenberg⁶ pointed out that in both Fermi's and Born's non-linear field theories the introduction of a non-linear term representing particle interactions brought with it the appearance of a fundamental length, and suggested that 'the introduction of a universal length may be bound up with a new fundamental modification of the formalism, just as c led to a modification of pre-relativity physics.' Again in 1938 Heisenberg⁸ expressed the view that a 'constant of the dimensions of a length (of order e^2/mc^2) plays a fundamental role in the theory of the elementary particles.' It is also relevant that Yukawa derived his prediction of the meson mass from a length appearing in a wave-equation for nuclear states.

The conclusions which may be drawn from these studies are that fundamental lengths are necessary to any theory of structure; that they are present in current quantum and nuclear theory, though at present playing a part secondary to that of dynamical analogies; and

¹ R. Fürth, *Zeits. f. Phys.*, 1929, **57**, 429

² Bohr, *J. Chem. Soc.*, 1932, p. 349

³ M. Born, *Proc. Roy. Soc. (A)*, 1933, **143**, 410

⁴ M. Born, *Proc. Ind. Acad. Sci. II*, 1935, **6**, 533

⁵ B. Podolsky, *Phys. Rev.*, 1934, **46**, 734

⁶ W. Heisenberg, *Zeits. f. Phys.*, 1936, **101**, 533

⁷ *Ibid.*, 1938, **110**, 251

⁸ Heisenberg, *Ann. d. Phys.*, 1938, **32**, 20

⁹ H. S. Flint, *Proc. Phys. Soc.*, 1936, **48**, 433

¹⁰ A. March, *Zeits. f. Phys.*, 1936, **104**, 93, 161

FUNDAMENTAL PHYSICAL THEORY

that they determine (a) particle interactions and transformations, and (b) the linear scale of all stationary states involved in fundamental structure.

In current theory the chief variable can be taken to be rest-mass rather than length (Heisenberg ^{1, 2}; Born ^{3, 4}), each particle requiring either a rest-mass M or the equivalent length given by $L = h/Mc$. But as Heisenberg pointed out, lengths are more appropriate, since they determine the particle transformations which occur. Moreover particle masses are inferences, dependent on dynamical conceptions, from spatial observations. Behind the speculations about 'minimum lengths,' 'finite size of the electron,' etc., lies the single incontrovertible fact of observation that a constant of length is indispensable to fundamental theory. Everything else rests on assumptions which may have to be discarded.

We shall therefore assume that the laws of the primary field involve one fundamental unit of length. This raises the question why current particle physics uses so many fundamental lengths. In 1898 the length appropriate to atomic structure was lacking; in 1950 the theory of particles uses at least *four* different combinations of constants providing fundamental lengths. We shall see that while at the opening of the century the need was to introduce a new fundamental length, the present task of theoretical research is to eliminate those which are redundant.

There is an antagonism between (dynamical or kinematic) dimensional constants involving both L and T , on the one hand, and simple (structural) constants of length, on the other hand. The gravitational constant, e/m , and c , could only be defined because in these three separate classical realms no characteristic lengths were present. But now that universal lengths have appeared in the more comprehensive quantum laws combining two of the classical realms, the earlier dynamical constants are losing something of their original significance, and ratios of lengths are becoming more important. Thus the history of physical theory during the last hundred years shows a movement from a period dominated by dimensional constants involving both L and T , through a transition stage using constants of both types and hence compelled to use complex approximative methods, towards—possibly—a non-dimensional formulation using

¹ Heisenberg, *Ann. d. Phys.*, 1938, **32**, 20

² Heisenberg, *Two Lectures*

³ M. Born, *Report Camb. Int. Conf. (Phys. Soc.)*, 1947, p. 16

⁴ M. Born, *Proc. Roy. Soc. Edin.*, 1949, **62**, 470

ratios of variable observed lengths to fundamental lengths, and a derived time parameter where appropriate. This corresponds to Heisenberg's suggestion that the correct method of introducing the universal length into physical theory may involve an entirely new formalism. The test of such a new method would be its power to account for the multiplicity of fundamental lengths in current four-coordinate theory.

4 *The Pure Number Constants of Atomic Theory*

The five primary dimensional constants of atomic theory (c , e , h , m , M) yield two non-dimensional combinations (α , β) with numerical values (1/137.03... and 1836...) known to at least three significant figures, which have no agreed theoretical significance. These values cannot be regarded as having no physical meaning, since they determine (a) the relative linear scales, and (b) the degree of interaction, of different types of elementary processes (nucleon, electronic, photon, etc.); and hence also (c) the scope and accuracy of the separate theories used for representing these processes. Thus a comprehensive theory, combining the special theories of electrons, photons, nucleons, mesons, etc., must account for α and β by providing a physically significant interpretation of their values which leads to an explanation of the origins of c , e , h , etc. in terms of some more general principle. As Bohr¹ has said, 'we must expect that the determination of these constants will be an integral part of a general consistent theory in which the existence of elementary electric particles and the existence of a quantum of action are both naturally incorporated.' A valid theory of these constants would provide more information about the special phenomena than is given by the earlier special theories. This means, for example, that a theory of α must describe the interactions of electrons and photons more accurately and simply than current theories.

The aim of a comprehensive derivation of α and β is the expression today of the unification of quantitative knowledge which has been the guiding principle in the advance of physical theory. α , β , and the other non-dimensional ratios (of meson masses, gravitational and cosmological constants, etc.) represent the only physical numbers yet determined by experiment whose values taken by themselves have any

¹ Bohr, *J. Chem. Soc.*, 1932, p. 349

FUNDAMENTAL PHYSICAL THEORY

physical significance, and that significance is still unknown. The challenge cannot be evaded.

It is probable that this task will require radically new ideas and the re-interpretation of a great part of physics. Moreover undue emphasis on the numerical values of α and β , separating them from the experimental background of the five dimensional constants from which they derive their importance, involves the risk of leading theoretical enquiry towards a mere numerology with no firm basis in lucid physical principles well rooted in empirical fact. As a result, in spite of the supreme importance of α and β , little of permanent value has yet been published on the problem of their derivation.

Yet the explanation of these numbers, which in 1930 seemed to many a fanciful idea, has now in 1950 become an urgent necessity for the advance of fundamental theory, and the interpretation of α is widely recognised as a major outstanding problem in fundamental physics. New experimental work on particle properties and theoretical investigations on particle interactions have both forced α to the front, and it is now clear that no valid synthesis of the present separate theories of material and light waves, and of the different types of particles, is possible without a significant derivation of α and β . In the following survey of work on these numbers emphasis will be put on α , on the assumption that it must be understood first.

In 1909 Einstein¹ stressed the importance of the fact that e^2/c has the same dimensions as h (though of the order $10^{-3}h$), and suggested that since both constants were foreign to the Maxwell-Lorentz theory, the future modification of that theory which leads to e must also account for h . This remark partially anticipated the aim of a unitary theory of material- and light-waves determining the value of α , put forward by Heisenberg in 1934.

In 1913 Jeans² pointed out that $hc/2\pi$ is almost equal to $(4\pi e)^2$ and asked 'is the new unit h anything more than a reappearance of the old unit $(4\pi e)$?' A year later³ he suggested that 'the atomicity of h may be associated with the atomicity of e . . . there is perhaps a hope that the two atomicities may be special aspects of some principle more general than either of them' (e.g. of a modification of the Maxwell equations).

¹ A. Einstein, *Phys. Zeits.*, 1909, **10**, 193

² J. H. Jeans, *Report of British Assoc.*, Birmingham, 1913, p. 380

³ J. H. Jeans, *Report on Radiation and Quantum Theory*, 1914, p. 81

L. L. WHYTE

Sommerfeld in 1916¹ called attention to the role of the non-dimensional combination, $2\pi e^2/hc$, appearing in his relativistic theory of the hydrogen spectrum fine-structure, and it thus acquired its name. It is of interest that α was first explicitly used in physical theory to represent ratios of *lengths*, i.e. of wave-lengths. Now it is known to determine also the hyper-fine structure of line spectra.

A step towards the understanding of α was taken by Perles,² who in effect combined the Einstein-Jeans idea with Bohr's view of h as introducing stability into certain electron orbits. Jeans had stressed *atomicity* as common to e and h ; Perles emphasised *stability*. He pointed out that e expressed the stability of a particular quantity of electric charge, h the stability of a particular angular momentum, and suggested that these two stabilities, being both in contradiction to classical theory, might be consequences of one more general principle.

This suggestion provides the most advanced interpretation of the common characteristics of e and h , and therefore of a possible common origin. Moreover it can be made more specific. If electronic charges and angular momenta were not quantised in discrete units, there would be no atomic stability and no measurable parameters characteristic of atomic structure. Thus e and h may be regarded as determining two kinds of stability, of which at least one must be present if consistent measurements of elementary processes are to be possible. e and h may both determine *measurable stability* in fundamental phenomena. This possibility of interpreting e and h , and therefore α , in terms of stability is the more interesting because β plays a fundamental role in the stability of nuclei.³ Moreover there has as yet been no really fundamental theory of stability using atomic ideas.

It is possible to carry this idea a stage further. For the stabilities which e and h represent must today be interpreted, not merely as those of single charges and orbits, but of all *nearly-stationary* states of simple or compound systems. (This non-relativistic term is used for all states which persist for a finite time, with either no changes at all, or only cyclic or reversible changes, to within some finite approximation. A nearly-stationary state may be taken, for present purposes, to be one of nearly constant energy with a sufficiently small exponential decay factor). It appears therefore that e and h may be associated with two

¹ A. Sommerfeld, *Ann. d. Phys.*, 1916, **51**, 51

² J. Perles, *Naturw.*, 1928, **16**, 1094

³ Bohr, *J. Chem. Soc.*, 1932, p. 349

FUNDAMENTAL PHYSICAL THEORY

expressions of the nearly-stationary property which must be present if consistent measurements of elementary processes are to be possible. In fact α may express the relation of two nearly-stationary measurable aspects of a single more general kind of process, or more precisely, the geometrical relation of two nearly-stationary, partially separable, asymptotic components of what we have called the primary field. α would then determine the relative magnitudes of, and degree of separability or interaction between, these asymptotic components.

On this view what has been regarded as 'atomicity' or 'discontinuous quantisation,' in relation to charge (e) and action or angular momentum (h), may be the expression of the fact that only certain unique components of the primary field of definite linear scale possess this nearly-stationary property. This interpretation is an advance, because 'atomicity' is an absolute, logically irreducible property, while nearly-stationary components are stationary only within a finite approximation and for a finite period of time, and are therefore open to experimental and theoretical investigation as components of more complex processes including exponential decay or growth factors.

Einstein, Jeans, and Perles made only tentative suggestions about a derivation of α . The first attempt to derive α or β was probably Reissner's,¹ who showed that the existence of the two elementary charges of opposite signs with different masses could be accounted for by adding a suitable term to the energy-function in his theory of point-charges. However this is now of little interest, since it was merely an attempt to improve classical electron theory.

In 1929 Eddington² published his well-known theory of α , relating it to the number of symmetrical terms in a basic probability matrix, and others made relatively trivial attempts to derive α or β . But it is now clear that no theory of α can be satisfactory which does not arise as part of a general theory of particle interactions and transformations. None of the theories of α published up to 1949 satisfy this requirement. Eddington's work on α is of the highest methodological interest, but fundamental particle physics has already moved into a new field on which his theory seems to throw little light.

A different approach to the interpretation of α was put forward by Whyte,^{3, 4} using the ideas discussed in section 3. As we have seen,

¹ H. Reissner, *Zeits. f. Phys.*, 1925, **31**, 844

² A. S. Eddington, *Proc. Phys. Soc. (A)*, 1929, **122**, 358

³ Whyte, *Critique of Physics*, p. 24

⁴ Whyte, *Zeits. f. Phys.*, 1929, **56**, 809 ; 1930, **61**, 274

with the introduction of h in addition to c , e , and m , fundamental theory had not only made good the earlier deficiency of scale, but had provided itself with a series of natural lengths, each appropriate to a different type of physical phenomenon, whose ratios involve α . This suggests that α has ultimately to be interpreted as the ratio of different lengths characteristic of various classes of elementary phenomena, and the problem of the duality of e and h (and of the value of α) solved by deriving all such lengths as functions of *one* fundamental length. In other words α is to be regarded as determining the relative spatial magnitude of the different phenomena involving the 'electron radius', the Compton wave-length, the radius of the first Bohr orbit, and the Rydberg wave-length respectively. This interpretation of α was proposed as part of a general research programme aimed at deriving α , β , and the ratio of electrostatic to gravitational forces as ratios of lengths appearing in the four-coordinate representation of a primary phenomenon whose laws do not involve x , y , z , and t as independent variables.

In the same year Fürth¹ also proposed that α and β should be regarded as ratios of lengths, in an attempt to derive β in terms of α .

In 1934 Podolsky² used the same interpretation of α as the ratio of apparently independent fundamental lengths, supported the suggestion that the primary role of h was to provide a universal length appropriate to atomic theory, and carried the problem a stage further. He sought to eliminate e and m from fundamental theory and to use instead c and two basic lengths, e^2/mc^2 and h/mc . The ultimate theory should be reached in *three* stages, providing first a derivation of α , then of β , and finally of the ratio of electrical to gravitational forces. Podolsky suggested that a valid derivation of α must require, *first* an explanation of the way in which electrons and positrons interact showing that the requirements of the theory can be satisfied if the interaction corresponds to $e^2 = hc \cdot \alpha / 2\pi$; *second*, an explanation of quantisation, showing that the theory requires $h = 2\pi e^2 / \alpha \cdot c$; and *third*, an explanation of the nature of radiation, showing why the limiting velocity of relativistic dynamics should be $c = 2\pi e^2 / \alpha \cdot h$.

Born³ in 1935 independently developed this interpretation in an essay devoted to α , stating that 'the explanation of (α) must be the

¹ Fürth, *Zeits. f. Phys.* 1929, **57**, 429

² Podolsky, *Phys. Rev.*, 1934, **46**, 734

³ Born, *Proc. Ind. Acad. Sci.* II, 1935, **6**, 533

FUNDAMENTAL PHYSICAL THEORY

central problem of natural philosophy,' that the existence of α can be ascribed to the fact that there are two different natural units of length in our laws (the quantum unit h/mc and the smaller unit e^2/mc^2), and that β 'determines the order of magnitude of all nuclear motions . . . and all properties of matter depending on these motions,' when expressed in units based on e , h , and m . He also made an attempt to derive α , which does not satisfy our present criteria.

In 1937 Flint¹ also suggested this interpretation of α . But twenty years have now passed since this was first put forward, without any understanding being gained of the manner in which several apparently independent lengths with ratios involving α appear in current theories of elementary processes. Zwicky² in 1937 proposed an investigation of all dimensionless ratios between significant physical quantities, but the main advance in this direction lies in the field of cosmology, where *lengths* do not play the same unique role.

Meanwhile the idea gradually developed that electrons and photons might be regarded as partially separable components of some single more general phenomenon, and that the error involved in treating them as separate was a function of α . Bohr pointed out that the error involved in the quantum-mechanical neglect of radiation-reaction was of the order α^2 , and Whyte³ suggested, more generally, that the inaccuracy in treating electrons and light as independent processes, rather than as two limiting aspects of a single process, was of this order. Here we see emerging the concept of a single comprehensive phenomenon, or primary field, of which particles such as electrons, photons, etc., represent special limiting components.

Heisenberg⁴ considered it improbable that a theory of α can be found until the short-range high-energy phenomena dependent on e^2/mc^2 are understood, and more recently⁵ he suggested that the value of α might be fixed by the condition that a primary formulation, as yet unknown, must correspond asymptotically (for larger distances) to quantum mechanics. Much recent work appears to reinforce Bohr's early suggestion⁶ that '*the whole attack on atomic problems leaning on the correspondence argument is an essentially approximative*

¹ H. S. Flint, *Proc. Roy. Soc. (A)*, 1937, **159**, 45

² F. Zwicky, *Proc. Nat. Acad. Sci.*, 1937, **23**, p. 106

³ Whyte, *Critique of Physics*, p. 106

⁴ Heisenberg, *Ann. d. Phys.*, 1938, **32**, 20

⁵ Heisenberg, *Two Lectures*

⁶ Bohr, *J. Chem. Soc.*, 1932, p. 349

procedure made possible only by the smallness of α' (my italics). It is clear that in the nucleus dynamical analogies do not even provide a useful first approximation.

We may summarise the result of this survey of published work on α thus :

The numerical values of α , β , and other pure-number ratios appearing in atomic physics are to be derived as an integral part of a general theory of the transformations of elementary parameters, from the condition of the asymptotic ($r \rightarrow \infty$) separability of particle-tracks and wave-fields as nearly-stationary components of a primary field whose mathematico-physical character has still to be identified. The crucial step in a derivation of α is to show how a single set of assumptions can account for the appearance of a plurality of fundamental lengths in current four-coordinate theory.

5 The Limits of Space-time Localisation

As physical techniques have become more exact and emphasis has passed from the approximate measurement of macroscopic phenomena to molecular and atomic properties, and finally to the nucleus, increasingly severe restrictions have been set to the use of space and time coordinates :

(a) As far back as 1930 Heisenberg wrote : ' In addition to the modifications of our ordinary space-time world required by the theory of relativity and characterised by the constant c , and to the inexactitude relations of the quantum theory symbolised by Planck's constant h , still other limitations will appear connected with the universal constants e , m , M . ' Today we can understand something of the further limitations set by the finitude of α and e^2/mc^2 , as the expression of the fact that even the test-objects used for measuring fields, and the measuring rods for determining lengths, have an ultimate structure which has to be taken into account. When this is done a new complementarity appears between the determination of exact field values and the coordinates of the test-body, and thus non-localisable fields have to be used (Markow, Heisenberg, Born, Yukawa, etc). But the full physical significance of this is not yet clear.

(b) The difficulty of meeting the requirements of relativity and quantum theory simultaneously, i.e. of expressing in invariant form the finite lengths necessary to a theory of structure (which parallels the difficulties of the relativistic dynamics of rigid bodies), suggests that

FUNDAMENTAL PHYSICAL THEORY

this may not be the most fertile line of advance. The separate limits set by c and by h to dynamical methods may correspond to two essentially distinct limiting cases in which some non-dynamical law (not using four-coordinates) can be given approximate dynamical representation. In other words the relativistic non-localisable field theory of the finite electron (e.g. see Markow¹) may represent a transition to some new formulation not based on dynamical analogies.

(c) In particular since 1927 it has become probable that within the nucleus, and in all regions of the order of e^2/mc^2 , it may be necessary to renounce exact space-time description and dynamical analogies. But many studies of this problem have been too naively classical. Classical methods cannot be used to determine precisely the character and extent of their own failure.

The net result of these theoretical studies and of recent experimental work bearing on this problem may be stated as follows. Some feature of current quantum-mechanical description based on dynamical analogies must be discarded in processes (i) within the nucleus; (ii) involving high energies in small regions of the order e^2/mc^2 ; or (iii) involving particle transformations. Four-coordinate representation appears to break down here.

But this does not necessarily mean that all quantitative coordinates must here be discarded, or that geometrical (three-dimensional) properties may not persist. It may be sufficient to give up the use of t as an independent variable associated with points in space. This is a reasonable hypothesis, in view of the passive role of the time coordinate in the stationary states of quantum mechanics and in the geometrical and wave optics of photons and electrons under conditions where transit-time effects can be neglected.²

The use of a time coordinate, associated with points throughout space, as an independent variable in partial differential equations is a highly sophisticated method ultimately requiring to justify it a detailed relativistic theory of the measurement of minute time intervals and of extended stationary wave-fields. But the use of a time coordinate might be replaced in fundamental theory (e.g. in small finite regions, or in relation to 'rigid' structures in the nucleus or elsewhere) by the use of a proper-time parameter associated with the processes of a

¹ Markow, *J. Phys. U.S.S.R.*, 1940, **2**, 453

² W. Ehrenberg and R. E. Siday, *Proc. Phys. Soc. (B)*, 1949, **62**, 8

finite system. Many such quantities, associated not with points in space, but with systems of finite size, are already in use : spatial size, total mass and energy, angular momentum, entropy, polarisation of structures, etc.,¹ and it may be necessary to add to these the proper-time of finite systems. This is one possible line of advance.

Another possibility has been opened up by Born's Principle of Reciprocity,² which asserts the invariance of fundamental laws under the exchanging of the space coordinates X and the corresponding momenta P . This cannot be discussed here. It has recently been taken up by Yukawa, and is connected with the work on non-localised fields.

Behind these recent studies there lies a possibility of great interest. It may be that physical theory can choose between a relativistic, four-coordinate, quasi-dynamical, contact-action, field representation of phenomena, and some other method more suited to a theory of fundamental structure. This other method may give priority of status to spatial relations, and use finite lengths, unique centres, and selected coordinate systems, thus emphasising the existence of unique points, lengths, and frames, e.g. in scattering processes (cf. Born³). The competition between these two tendencies is seen in Schrödinger's equation for the hydrogen atom, for here in a field representation using partial differential equations and the idea of contact action there is the stubborn kernel of the Coulomb term emphasising the importance of finite distances. The coming years may see this conflict between contact action and action at a distance, and between field and matter, finally resolved.

6 Conclusion

I It is possible that the next major advance in fundamental theory will carry physical theory beyond dynamical analogies to an entirely new method more appropriate to the representation of structure. At such a time theoretical enquiries have a special value, yet reliable arguments cannot be based on established methods, as these are known to be inadequate. Investigations must therefore be guided by relying as far as possible on the most directly observable quantities, i.e. those involving the fewest theoretical assumptions.

¹ Cf. P. A. M. Dirac, *Phys. Rev.*, 1948, **73**, 1092

² M. Born, for references see *Nature*, 1950, **165**, 270

³ Born, *Report Camb. Int. Conf. (Phys. Soc.)*, 1947, p. 14

FUNDAMENTAL PHYSICAL THEORY

II If this view is correct the next advance may involve :

<i>the elimination from fundamental theory of:</i>	<i>and the substitution of:</i>
Invariant four - coordinate representation and the use of t as an independent variable.	Finite three-dimensional spatial relations, with a derived proper time associated with finite systems.
Elementary particles. Wave-particle duality.	Sets of elementary parameters. Parametric structure of a phenomenon.
Action and interaction by contact or at a distance. Forces between two entities.	Changing structure of finite, complex systems. n -term changing spatial relations.
Quantisation of energy and action.	Definiteness of scale of changing patterns of nucleons.

III The key to further advance may lie in the interpretation of the numerical value of α as resulting from the representation of a primary field, not involving an independent time coordinate, in terms of four coordinates. Four-coordinate physics has established a series of approximations to the primary field, depending in most cases on the smallness of α , i.e. on the smallness of certain ratios of natural lengths appearing in coordinate theory. But the formulation of the laws holding in the nucleus, in small regions, and in particle transformations involves the elimination of α , using a new method of representation appropriate to scale-fixed phenomena.

IV This interpretation of the position of particle physics carries the following implications for other branches of physics :

Low temperature physics may require a theory based on a new concept of structure, rather than on dynamical analogies, since at very low temperatures the effective range of nuclear (structural) laws may extend to macroscopic regions.

Cosmological theories involving a high extrapolation of atomic principles may require revision if the fundamental concepts of atomic physics undergo a radical transformation.

L. L. WHYTE

REVIEWS

The Concept of Mind, Gilbert Ryle, Hutchinson's University Library, London, 1949. Pp. 334. 12s. 6d.

PROFESSOR RYLE'S aim in writing this book was not to provide new information about minds, but to 'rectify the local geography of the knowledge which we already possess.' A philosopher like Hobbes assumed that it is proper to speak of the facts of mental life in terms of the categories of physical science. But to speak in this way is to allocate the facts of mental life to one 'logical type or category' when they really belong to another. Descartes, on the other hand, being convinced of the inapplicability to mental life of the mechanistic hypothesis which he applied in the non-mental sphere, postulated an invisible existent called 'mind,' mysteriously situated in a body governed by mechanical laws. According to this myth of 'the ghost in the machine,' which makes the mind an invisible 'thing' or 'substance,' invisible events and operations stand in a causal relation to observable events and activities. It is this 'para-mechanical theory' which Professor Ryle is concerned to destroy. What he wishes to show is not that there is no mental life but that, in the Cartesian myth, the facts of mental life are allocated to one range of logical types or categories when they actually belong to another. For example, mind is spoken of as a 'thing' or 'substance' of which it is proper to say that it is situated 'in.' The Cartesian myth thus remedies the Hobbist myth only by duplicating it.

As far as the 'ghost in the machine' is concerned, I think that Professor Ryle disposes of it in a telling way. And he rightly points out that, when this 'myth' has been disposed of, the considerable difficulty which some modern philosophers profess to find in discovering how we know other minds should be disposed of also. But I do not think that it requires any very profound reflection on the content of the book in order to see that it is not simply Descartes' 'ghost in the machine' which suffers from the author's attack. Take the following sentence, for example. 'To find that most people have minds (though idiots and infants in arms do not) is simply to find that they are all able and prone to do certain sorts of things, and this we do by witnessing the sort of things they do.' Professor Ryle, does not, of course, question the propriety of speaking about 'minds'; but the sentence just quoted seems to state pretty clearly that the word 'mind' means a complex of abilities and dispositions. To think that the word means a spiritual 'substance' or 'form' is to behave like a foreign visitor to Oxford who thinks that 'the university' is a member of the class of which St John's College, the Bodleian, etc., are members and who accordingly, having seen the various buildings which visitors ought to see,

REVIEWS

asks to be shown 'the university.' But, if the man who thinks of the mind as something over and above a complex of abilities and dispositions commits an analogous category-mistake, it is obviously more than the Cartesian theory of 'the ghost in the machine' which is affected. It might be said that to admit certain dispositions is to admit some spiritual entities; but Professor Ryle will not allow that dispositional statements describe existents: 'if they are true, they are satisfied by narrated incidents.'

Now, the answer to the question whether the mind (and I had better make it clear that by 'mind' the author does not mean simply 'the intellect' but rather what many people would speak of as 'the soul') can be properly described in the way that Professor Ryle describes it obviously depends on the answer to the question whether there is any good evidence which warrants or necessitates our describing it in any other way. And the author argues that there is no such evidence. But it is, I think, worth pointing out that an unwary reader might easily be led into misunderstanding the nature of the problem. For one thing, the use of phrases like 'category-mistake,' 'allocation of concepts to logical types,' 'the logic of the problem,' might suggest to him that the problem of mind is a logical problem. But, though one can doubtless speak illogically about minds, the problem of the nature of mind is not a problem of logic, in the normal sense of 'logic' and 'logical problem.' The author would not, of course, maintain that the problem of mind is a logical problem in the sense of being a problem of formal logic; but the reader might be led into thinking that all metaphysical theories of the mind or soul are crimes against logic, without realising that the question whether this is so or not cannot be settled by logic itself. Again, as Professor Ryle is constantly analysing concepts (a very proper procedure, of course) a reader might receive the impression that an analysis of ordinary language disposes, without more ado, of metaphysical theories of the mind. But 'ordinary people' (and not only 'ordinary people') frequently cannot give any clear account of what they mean when they make statements about minds; and, in addition, what they do mean is often, sometimes unconsciously, determined, or at least influenced, by preconceptions of one kind or another. Analysis of meaning thus inevitably tends to become an analysis of what people *ought* to mean by their statements, that is, if they attended to the facts of evidence. Thus to say that Cartesians and Platonists (and presumably also Thomists and even Kantians) allocate mental concepts to wrong logical types and that their theories are not supported by an analysis of ordinary language is in the long run, I think, equivalent to saying that there are no facts of mental life which can properly be adduced in support of their theories.

Professor Ryle is eminently a common-sense philosopher, with a strong dislike for pomposity and appeals to privileged experiences. This non-sense attitude is at once a strength and a weakness. Its strength comes

REVIEWS

out, for example, in the author's discussion of the sense datum theory. When he maintains that a man observing a horse-race is observing a horse-race and not intuiting a patchwork of colours, he is talking common sense ; and his contention that the sense datum theorists confuse concepts of sensation with concepts of observation is well supported. On the other hand, the weakness of the no-nonsense attitude is revealed, I think, both positively and negatively ; positively in the discussion of certain topics and negatively in the omission of certain themes which might well be relevant to the concept of mind.

In the chapter on self-knowledge there is a section on the self, or the notion of 'I.' Now Professor Ryle admits that 'even Hume confesses that, when he has tried to sketch all the items of his experience, he has found nothing there to answer to the word "I," and yet he is not satisfied that there does not remain something more and something important, without which his sketch fails to describe his experience.' But Professor Ryle does not seem to be troubled by Hume's after-thoughts. 'Gratuitous mystification begins from the moment that we start to peer around for the beings named by our pronouns.' Well, mystification does sometimes result from such peering ; we have only to look at the antics that Fichte performs when he is chasing the 'I.' But is the peering gratuitous ? The author admits (he could hardly do otherwise) that sentences containing pronouns do mention identifiable people. Now, if Professor Ryle said that the pronoun 'I' always means 'my body,' one might disagree but one would at least have a definite, if very naive, position to deal with. But he says nothing of the kind. On the contrary, he very rightly states that there are sentences where the pronoun 'I' could not be replaced by the words 'my body' without making nonsense. What, then, does 'I' mean when it cannot be replaced by 'my body' ? In particular, what of the 'I' and 'me' involved in self-consciousness ? Among other things the author points out that 'an ordinary reviewer may review a book, while a second order reviewer criticises reviews of the book. But the second order review is not a criticism of itself. It can only be criticised in a further third order review.' Similarly, I can describe my past self ; but I cannot get hold, so to speak, of my present self. Hence the feeling of the mysterious elusiveness of the 'I.' But, even if this is true, I fail to see that everything follows in regard to the nature of the 'I.' One gets the impression that the author is trying to clear up a linguistic puzzle by ingenious analogies ; and I do not think that the nature of the self is a linguistic puzzle which can be cleared up in this way. It is all very well to say, as the author does, that 'at a certain stage the child discovers the trick of directing higher order acts upon his own lower order acts' ; but this does not clear up the nature of the 'I.' It is always possible to ask what renders this sort of 'trick' possible. In other words, I am inclined to think that Professor Ryle evades the issue.

REVIEWS

I do not want to say much about mental activities and experiences which Professor Ryle omits to discuss, because he may have had good reasons for omitting any discussion of them. But, while one cannot justifiably dictate to an author what he is to include in his book, one is justified in pointing out that a relevant topic has been omitted, if one thinks that it has been omitted. Kant, for example, thought that the moral consciousness and moral experience (I fear that Professor Ryle would think these phrases excessively pompous) are relevant to the problem of the nature of the self; and on this point I agree with him. Professor Ryle does indeed devote a paragraph to the subject of conscience and he makes it clear that he regards it as one more example of an ability or inclination. For he speaks of the child who having listened to the didactic discourses of his parents and schoolmasters, 'has acquired some capacity and inclination to deliver refresher-lessons to himself in their magisterial tones of voice.' But the relevant point, I think, is not so much from what source one derives one's initial moral ideas as of what nature the 'I' is which, so it seems to me, is very much in evidence (and not at all a 'ghost in a machine') in important moral decisions. It appears to me that the 'I' is constantly slipping through the web of analysis woven by the author. I suppose that Professor Ryle would say that, if I mean by the 'I' an immaterial 'substance' or 'form,' it cannot be caught in the web because it is not there to be caught. He is not out to catch ghosts. And, if I reply that I mean by the 'I' the whole man rather than simply an immaterial 'mind,' he might say, 'Precisely; so do I; in some sentences at least.' But, as Professor Ryle freely admits that a man is a man, and not simply an animal, what is it that peculiarly distinguishes him as such? Certain abilities and dispositions? Yes; but are these separate entities or not? If not, do they belong to the body or not? If not, to what? Obviously we are back at the beginning again; for Professor Ryle would doubtless say that my language is all wrong and that I am committing category-mistakes which cannot be avoided unless I drop the hypothesis which leads to them. The trouble is, of course, that one has to use a language, the range of which is strictly limited. We talk about 'clear minds'; but we do not confuse them with clear soup. And how could the theory of an immaterial mind ever have been thought of, were Professor Ryle's account of mind an adequate account? It may or may not involve category-mistakes; but it is not itself, I think, the result of a category-mistake. It seems to me that moral and religious 'experiences' are relevant to its origin.

In any case, reflecting on what Professor Ryle has to say is something worth doing. A review cannot do justice to his book, which really deserves a commentary. It is an important book, dealing with an important subject; and it is a very clearly written book, in the best British tradition of philosophical writing. It is, moreover, enriched by humour and by the frequent

REVIEWS

use of (often homely) illustration and analogy. A reviewer who has more sympathy for metaphysics than appears to be shown by the Waynflete Professor of Metaphysical Philosophy at Oxford will naturally tend to draw attention to points with which he disagrees. But this does not mean that there may not be (and in this case there certainly is) a great deal with which he agrees. To take a minor topic as an example, I wish that all those who talk glibly about the 'subjectivity' of the secondary qualities would digest what Professor Ryle has to say on the matter. Moreover, even if I am not at one with the author in his concept of mind, I know of no other work which gives so able and careful and acute an exposition and defence of the sort of view he advances. I feel confident that the book, which covers such a wide range and contains analysis of so many departments of mental life, will prove to be of a much more than passing interest and importance. Certainly, no British philosopher, whatever his 'colour,' can afford to neglect it. One does not need to agree in all things with an author in order to realise that his book is a work of major importance.

FREDERICK C. COPLESTON

The Origins of Modern Science, Herbert Butterfield, G. Bell and Sons, London, 1950. Pp. x + 217. 10s. 6d.

ANYONE who does not regard scientific theories as fixed dogmas will derive much pleasure from reading Professor Butterfield's book. One of his principal themes, if I have not misunderstood him, is that a given scientific theory, at a given epoch, represents the best that contemporary intelligence can do with a particular problem on the basis of the beliefs which they have received from their predecessors, that this theory then fixes peoples' ideas and prevents their successors from viewing the problem in any other light, until someone arrives on the scene who is able to perform the exceedingly difficult task of 'changing his mind.' Usually he does not succeed in changing the minds of his contemporaries, but gradually the changed outlook spreads and becomes a new orthodoxy. As the author says, 'It will concern us particularly to take note of those cases in which men not only solved a problem but had to alter their mentality in the process, or at least discovered afterwards that the solution involved a change in their mental approach.' He also says that the subject has not been turned into genuine history if we 'remain content with a merely biographical mode of treatment, and particularly if we construct our story of science by drawing straight lines from one great figure to another. . . . It has proved almost more useful to learn something of the misfires and the mistaken hypotheses of early scientists. . . . And it is necessary on each occasion to have a picture of the older systems—the type of science that was overthrown' (p. ix).

Another useful thing that Professor Butterfield does is to challenge the

REVIEWS

doctrine, which many historians of science have taught, that the revolution in science (as he calls it) was the result of the discovery of the method of observation and experiment. The important part played by synthesisers and mathematicians is fully illustrated. The story of Galileo dropping cannon balls from the top of the tower at Pisa (on which we are all brought up) is exploded on page 70. 'To crown the comedy, it was an Aristotelian, Coresio, who in 1612 claimed that previous experiments had been carried out from too low an altitude. In a work published in that year he described how he had improved on all previous attempts—he had not merely dropped bodies from a high window, he had gone to the very top of the tower of Pisa.' Professor Butterfield also points out that in one of the dialogues of Galileo 'it is Simplicius, the spokesman of the Aristotelians—the butt of the whole piece—who defends the experimental method of Aristotle against what is described as the mathematical method of Galileo.' Boyle, who was a great experimenter and professed to follow Bacon in disdaining system-building (although he was at the same time a fervent believer in the corpuscular philosophy) did not revolutionise chemistry to the extent achieved by Lavoisier who 'was not one of those men who are ingenious in experimental devices, but he seized upon the work of his contemporaries and the hints that were scattered over a century of chemical history, and used them to some purpose' (p. 188).

The title does not very well suggest the scope of the book, which deals with development as well as origins. Moreover, as the author says, 'The history of science ought not merely to exist by itself in a separate pocket' but should attempt to show the significance of the intellectual changes for general history in its broadest sense. Two chapters are devoted to this topic.

There are one or two curiosities. On page 45 Harvey's book is referred to as *De motu cordium* instead of the more usual *cordis*. On page 205 a hyphen has exchanged lines with a 'd.' On page 167 the somewhat sinister-sounding phrase 'the dialectic of history' flashes upon the scene like a meteor and does not reappear. John Locke is mentioned several times, but David Hume not at all. Is this because his influence is not yet history?

By concentrating on essentials Professor Butterfield has achieved a great deal in his 217 pages which can be warmly recommended to readers of this *Journal*.

J. H. WOODGER

Definition, Richard Robinson, Clarendon Press, Oxford, 1950. Pp. 207. 15s.

MR ROBINSON's essay will be welcome not only to the philosopher of science but also to workers in various other fields of science and scholarship. For

the number of more or less comprehensive treatises on this subject is not large, and may be called very small indeed if compared with modern literature on deduction. This is a striking phenomenon, as definitions play a role in every domain of science and are also important for practice, which cannot be said of deduction. On the other hand, the variety of its fields of application makes definition a very difficult subject.

These considerations might induce us not to apply an exacting standard to Robinson's performance ; but this would be out of the question in view of the importance of the theory of definition and moreover it would not do justice to the many qualities of the book.

Before going into details, I may be allowed to sum up my appreciation as follows. Robinson's book gives provocative and stimulating analyses of nearly every important problem in the theory of definition and therefore constitutes a precious contribution to the literature on this topic ; it does not, however, provide the reader with a more or less decisive solution of these problems.

A student of modern logic will sometimes be surprised by a certain neglect of the formal side of the problems. For instance, on page 15 a fundamental classification of definitions is based on a distinction between the *purpose* and the *method* of a definition. However, it seems inevitable to consider also the manner in which a definition is stated. The problem, raised in ch. I, § 7, as to what sort of entity a definition does apply to, appears to be clarified in many respects if approached from the formal side. For the opinions mentioned there seem to point to the possibility of stating materially the same definition in formally different manners. Another, related, question refers to the behaviour of definitions under translation.

Logicians may be inclined to suspect some connection between this neglect of the formal aspect and the outspoken acceptance of a psychological point of view. I feel that there is indeed such a connection, without blaming Robinson for adopting a psychological point of view which after all plays only a very subordinate role in his book. For as soon as we cease to restrict ourselves to the study of deduction and of deductive science, and extend our interest to other fields of thought and life, it seems inevitable sometimes to rely on psychological methods.

Nevertheless, we touch here what really is a weak point in Robinson's endeavour. For he seems not to realise that the problems of definition take quite different shapes when considered in relation to deductive science and in relation to other fields of thought. In fact, he first develops a general theory of definition, and then applies it to mathematics. And in mathematics, he fails to stress the fundamental distinction between the procedures applied in mathematical theories *in statu nascendi* and those applied in finished mathematical theories. In a systematic study the reverse order might be preferable. We could start from the study of definitions found in finished

REVIEWS

mathematical theories ; these definitions could be dealt with by applying the methods of modern logic. Then we could turn to definitions in mathematical theories *in statu nascendi* ; it would perhaps be most profitable to consider the earlier stages in the development of those mathematical theories which already were studied in their contemporary form. Finally we could explore the remaining fields of thought.

It remains to be seen whether such an approach, though more systematic, would turn out to be more fertile. For the role of definitions in deductive sciences seems to be entirely different from their function elsewhere. Logical methods are adequate in the study of deductive systems ; it may be questioned whether they will be of much use elsewhere. So it may well turn out to be impossible essentially to improve upon Robinson's methods and results.

The historical material scattered throughout the book certainly adds to its value. The influence of Locke on Robinson's views is conspicuous ; but it seems curious that he should not have consulted Leibniz's *Nouveaux Essais*.

EVERT W. BETH

Makers of Mathematics, Alfred Hooper, Faber and Faber, Ltd., London, 1949.
Pp. ix + 402. 18s.

THE examination of the concepts and procedures which are implicitly assumed in elementary mathematics is of great importance in coming to any appreciation of the methods and philosophy of mathematical enquiry and of the natural sciences. The importance of such studies was shown by the work of Felix Klein. To give some account of such matters is one approach that can be used in a 'popular' book on mathematics for the adult reader. Another approach is to describe the concepts which are at present exercising the minds of mathematicians, and to examine the comparatively simple ideas and, sometimes, puzzles in which they have originated.

The present book takes neither of these useful standpoints. It describes the ordinary material of elementary mathematics—from the idea of natural numbers to the infinitesimal calculus—against a background of the stories of the men with whose names various of these ideas and processes are associated in popular tradition. Unfortunately, the personal histories consist largely of anecdotes ; and the mathematical discoveries appear scattered in this background as a collection of more or less isolated facts. To the present reviewer the book gave no idea of the methods of mathematics, the significance and nature of the concepts it uses, the elegance of its results, or the majestic beauty of the whole canvas which has been painted by the great mathematicians.

REVIEWS

This is an English edition printed in this country : it is a pity that English usage has not been adopted in the spelling for this edition (for example, 'traveled' on page 274 and 'meager' on page 350). On page 222, the author speaks of 'odd prime numbers'—but all prime numbers are necessarily odd (although all odd numbers are not prime). Furthermore, the theorem quoted on this page and ascribed to Fermat is true of *all odd numbers* and not merely (as stated by the author) of those odd numbers that are prime. On page 130, 'similar' is used in error for 'same.' The description of Pascal's contact with the Jansenists (pages 226–232) appears somewhat superficial. It shows no acquaintance with an authoritative life of Pascal (as was given by Denzil Patrick) and little appreciation of Pascal's own assessment of the relative importance of his different activities. A greater acquaintance with the achievements of Greek mathematics would have saved the author his excessive praise of Descartes (page 209) as a mathematician.

A. R. MILLER

The Natural Philosophy of Plant Form, Agnes Arber, Cambridge University Press, 1950. Pp. xiv + 247. 25s.

The Natural Philosophy of Plant Form is a title to arrest the eye and stimulate the imagination ; it has a pleasing eighteenth-century ring about it in an age of atomic physics and plant physiology. Can it be that to understand the form of plants we have, after all, only to study their structure, read contentedly our Aristotle and Goethe and meditate upon the ultimate nature of things ? But as soon as we look into Mrs Arber's book this agreeable vision is dispelled ; the philosophical position is reached only after much difficulty : the author was led to it after applying to her work 'more universal, more stringent modes of thought' than those of analytical science. It so happens that these modes of thought are very much coloured by Greek philosophy and Mrs Arber's book ends more or less where it begins, in Athens. But in our excursion we enjoy a wealth of erudition, historical, philosophical and botanical, at which most botanists will marvel and feel very humble, yet also, nevertheless, very restless.

The early chapters introduce us to the botanical writings of a number of outstanding figures of whom Theophrastus, Albertus Magnus and, inevitably, Goethe, may be mentioned. These chapters are valuable in content and delightful to read ; they also indicate very clearly the influence the past has had on the development of Mrs Arber's thought. In the first chapter the meaning of the term 'morphology' is discussed and we have a definition of morphology breath-taking in its scope. 'It is hence the business of morphology to connect into one coherent whole all that may be held

REVIEWS

to belong to the intrinsic nature of a living being.' This is very sound, but is it surprising that, having dug a pit of these dimensions, Mrs Arber later on falls into it? It really will not do to say that experience has shown that a study of form 'often serves as an index to more recondite characters.' It can also be very misleading, as the existence of much pseudo-phylogeny based upon the study of living plants alone clearly shows. The genetic structure and thus, inevitably, its ancestry must belong to the 'intrinsic nature' of any organism, and a morphology which ignores phylogeny is incomplete from the beginning. Phylogeny receives little mention by Mrs Arber: it is disregarded, one suspects, because of the excessive imagination and speculation of many morphologists of the last century, but this is no sound reason for omitting it from a morphology which purports to be natural. A knowledge of palaeobotany and of the lower plants is essential in the understanding of plant form: they are not merely accessories to be welcomed when they support a conclusion of formal morphology, such as the partial-shoot theory of the leaf.

Mrs Arber's morphology, lacking the discipline of palaeobotany and genetics, is capable of almost anything because it is largely subjective. What could be more instructive than the mention of the controversy over the morphological interpretation of the female cone of the Conifers? For generations the battle has raged between the holders of one theory and those of another until very recently a Swedish botanist had the happy notion of studying in detail the fossil history of the group. As a result the interpretation of the female cone is no longer a matter of doubt or dispute. Beside these facts, which seem to have been too late for Mrs Arber to see, her additional theory is an unnecessary, albeit ingenious, complication. But happy in the task of working out a morphological system for the higher plants cut off from reality, Mrs Arber throws caution to the winds and produces a new and, according to her system, plausible theory of the morphological nature of the axillary bud. This is not the place to enter botanical objections, only to ask is it consistent with all the evidence, ought it not to be reconsidered in the light of 'a more universal, and also more stringent,' botany?

Mrs Arber has little to say about homology except to speak of 'members which are not homologous in the phylogenetic sense.' But can structures ever be homologous in any other sense if the concept of homology is to serve a useful purpose? We can, of course, have our private systems of morphology, if we will, and define homology in whatever way we wish, but our contribution to our fellow scientists' understanding of plant form will not be appreciable. In her discussion of the 'type' concept and of Troll's *Gestalttypus* Mrs Arber has many useful things to say. But no satisfactory argument is raised, or ever has been raised, against the view that the 'type'-effect is merely psychological. Does not the human mind tend to relate

REVIEWS

a group of similar structures and extract the essentials, the 'type'? Any other object possessing these essentials is then associated with that class. Mrs Arber's concept of 'parallel becoming' does not seem to liberate one from the tyranny of typology, it merely pushes the 'type' a little further back into obscurity. In a discussion of causation in plant morphology in the last chapter, the suggestion is made that the 'type' is an expression of the formal cause in the Aristotelian sense. But, since the formal cause is defined as 'the essence or essential nature of the thing' is this suggestion anything more than a tautology?

The last chapter is one of the most interesting and the whole treatment of causation stimulates one to argument. Mrs Arber performs a service in forcing the question before one of whether Aristotle's formal and final causes have any meaning in biology. Mrs Arber's claim that the study of the material and efficient causes, which are equated with the physico-chemical approach to biology, is compatible with that of the formal and final causes equated with the teleological approach, 'within the larger realm of synthetic biology,' might be conceded if one knew what this synthetic biology was. Does it really exist or is it just an unhappy marriage of impossible partners? If one is to use Aristotelian terminology, would it not be truer to say that natural selection provides the expression of a final cause, and the fact of genetical determination (that an organism is already determined in its fertilised gamete) that of the formal cause? This way of thinking may also lead to a better understanding of the 'type.' It is very probable upon genetic grounds, as Whitehouse has shown, that Angiosperms are monophyletic. A basic set of genes may therefore be responsible for carpel formation and the occurrence of this organ in one form or another, but always recognisable, throughout the Angiosperms is explicable in a way that demands the study, not of types, but of genic and molecular symmetry.

Mrs Arber has spoken elsewhere of the difficulty of finding words to express morphological ideas, but some phrases in her present book cannot pass without comment. It is difficult to imagine the agitation of a vascular bundle 'striving after the stelar condition' or of organs showing 'the tendency of lateral structures to aspire to take up the role of the parent,' yet such expressions demand an emotional environment if they are to have meaning. They also inevitably bring in a suggestion of volition which is almost certainly not intended.

It is difficult to sum up a review of a book as novel as this. One thing is certain, we must be very grateful to Mrs Arber for having written it and for having produced such a feast of scholarship and stimulating ideas. It is a courageous and very fine attempt to re-establish an approach to plant morphology which finds, amongst those who think about morphology at all, very little favour at the present time. Whether or not there will be a swing from the physico-chemical to the teleological approach to the study

REVIEWS

of form, as the author appears to envisage, remains to be seen, but there is no doubt that to the future historians of botany the publication of *The Natural Philosophy of Plant Form* will be an important and significant event.

PETER BELL

SHORT NOTICE

The Philosophical Quarterly. Vol. I, No. 1, October, 1950. Pp. 96.
(Published by the University of St. Andrews for the Scots Philosophical Club.) Price 6s.

THIS new philosophical journal has been launched by the Scots Philosophical Club and is edited by Professor T. M. Knox of St. Andrews University. Its programme is to give preference to 'work on metaphysics, ethics, political theory and the philosophy of art, history and religion.' This is important and essential work which, in the current preoccupation with language, logic and method, is not adequately represented in English philosophical journals. The scientist who wishes to get some general picture of what philosophers are doing in the most diverse fields is likely to find it in these pages.

In introducing the journal to its readers Professor Kemp Smith, doyen of Scottish philosophers, reminds the reader that *Mind*, the leading philosophical journal in England, was in origin a Scottish venture. There could be no happier augury for the success of the new journal.

BOOKS RECEIVED FOR REVIEW

Inclusion of books in this list does not preclude their being reviewed in later issues

- Angus Armitage, *Sun, Stand Thou Still*, Sigma Books Ltd., London, 1947, pp. x + 210, 12s. 6d.
- E. W. Beth, *Les Fondements Logiques des Mathématiques*, Monographies réunies par Mme P. Destouches-Février, Paris, 1950, pp. 222.
- Frederick C. Copleston, S.J., *A History of Philosophy*, Vol. 2, 'Augustine to Scotus,' Bellarmine Series, Burns Oates & Washbourne Ltd., London, 1950, pp. x + 614, 25s.
- Maurice Cornforth, *In Defence of Philosophy*, Lawrence & Wishart, Ltd., London, 1950, pp. xv + 260, 12s. 6d.
- R. A. Fisher, *Creative Aspects of Natural Law*, Cambridge University Press, London, 1950, pp. v + 23, 2s.

REVIEWS

- Bruno v. Freytag gen. Loeringhoff, *Gedanken zur Philosophie der Mathematik*, Westkulturverlag Anton Hain, Meisenheim-am-Glan, 1948, pp. 56, DM 2.80.
- Eduard May, *Kleiner Grundriss der Naturphilosophie*, Westkulturverlag Anton Hain, Meisenheim-on-Glan, 1949, pp. 106, DM 4.90.
- Erwin Schrödinger, *Space-Time Structure*, Cambridge University Press, London, 1950, pp. viii + 119, 12s. 6d.
- Dorothy Stimson, *Scientists and Amateurs*, A History of the Royal Society, Sigma Books, Ltd., London, 1949, pp. xvi + 270, 15s.
- Symposia of the Society for Experimental Biology*, No. IV, *Physiological Mechanisms in Animal Behaviour*, Cambridge University Press, London, 1950, pp. vii + 482, 35s.
- W. P. D. Wightman, *The Growth of Scientific Ideas*, Oliver & Boyd, Edinburgh and London, 1950, pp. xii + 495, 25s.

MEETINGS OF THE PHILOSOPHY OF SCIENCE GROUP

The following is the programme arranged for this academic year. The meetings are held at 5.30 p.m. in the Joint Common Room, University College, Gower Street, London, W.C.1.

Monday, 16 October 1950 : Dr Philipp Frank on 'The Positivistic, the Meta-physical, and the Sociological Interpretations of Science'

Monday, 13 November 1950 : Mr D. M. Mackay on 'Cybernetics (Mind-like Behaviour in Artefacts)'

Monday, 11 December 1950 : Professor J. B. S. Haldane, F.R.S., on 'Inverse Probability'

Monday, 22 January 1951 : Professor J. Z. Young, F.R.S., on 'Biological Relativity'

Friday, 2 March 1951 : Professor E. W. Beth on 'The Existence of Mathematical Entities'

Monday, 5 March 1951 : ANNUAL GENERAL MEETING. The Chairman (Professor J. H. Woodger) will deliver an address on 'Science without Properties'.

Monday, 23 April 1951 : Dr N. W. Pirie on 'Concepts out of Context : The Pied Pipers of Science'

Monday, 21 May 1951 : Professor K. R. Popper on 'A Critical Discussion of Operationalism and Related Views'

Monday, 11 June 1951 : Mr L. L. Whyte on 'The Philosophy of Science : Today and Tomorrow'